DOCUMENT RESUME

ED 109 651 CS 002 056

AUTHOR Murphy, Richard T.
TITLE Adult Functional Reading Study; Project 1: Targeted Research & Development Reading Program Objective, Sub-parts 1, 2, and 3. Supplement to Final Report.
INSTITUTION Educational Testing Service, Princeton, N.J.
SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C.
REPORT NO ETS-PR-75-2
BUREAU NO BR-0-9004
PUB DATE Jan 75
CONTRACT OEC-0-70-4791(508)
NOTE 124p.; See related document CS 002 055; Page 11 has poor reproducibility because of type size.

EDRS PRICE MF-$0.76 HC-$5.70 PLUS POSTAGE
DESCRIPTORS *Adult Literacy; Cloze Procedure; Criterion Referenced Tests; *Functional Reading; High School Students; National Surveys; *Reading Research; *Reading Tests; Word Recognition
IDENTIFIERS *Adult Functional Reading Study

ABSTRACT

Articles in this report are based on the data obtained by the Adult Functional Reading Study. This study examined the reading activities of American adults and tested adults' performance on functional reading tasks. The supplement contains several articles concerned with the reading skills required to read and answer or perform the reading tasks and an analysis of reading competence and schooling related to economic benefits. Chapter topics are error analysis and inquiry, relationship of decoding to adult functional reading, relationship of functional reading to performance on cloze tests, basic reading competency in the schools, and the economic benefits of schooling and reading competence. (MKM)

SUPPLEMENT

to

FINAL REPORT

Project No. 0-9004

Grant No. OEC-0-70-4791 (508)

PR 75-2

ADULT FUNCTIONAL READING STUDY

Project I: Targeted Research & Development Reading Program Objective, Sub-parts 1, 2, and 3

Richard T. Murphy
Educational Testing Service
Princeton, New Jersey 08540

January, 1975

The research reported herein was performed pursuant to a contract with the National Institute of Education, U. S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official National Institute of Education position or policy.

U.S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

National Institute of Education

CONTENTS

Acknowledgements 1

Chapter 1 - ERROR ANALYSIS AND INQUIRY 5

Chapter 2 - RELATIONSHIP OF DECODING TO ADULT FUNCTIONAL READING by Carolyn Massad 17

References 43

Appendix 45

Chapter 3 - RELATIONSHIP OF FUNCTIONAL READING TO PERFORMANCE ON CLOZE TESTS by Spencer Swinton and Michael Zieky. 51

References 65

Chapter 4 - BASIC READING COMPETENCY IN THE SCHOOLS 67

Chapter 5 - THE ECONOMIC BENEFITS OF SCHOOLING AND READING COMPETENCE by Kan-Hua Young and Dean Jamison 73

Footnotes. 95

References 99

Figure and Tables. 103

Appendix A 113

Appendix B 119

Appendix C 123

Appendix D 127

ACKNOWLEDGEMENTS

During the period July 1970 to December 1973, the reading activities of American adults were surveyed; reading tasks, intended to be representative of the kinds of reading performed by adults in the course of a normal day, were developed and administered to a national probability sample of adults; and results of the surveys were reported in Reading Activities of American Adults (Sharon, 1972) and Adult Functional Reading Study (Murphy, 1973). As a supplement to the original study, the tasks listed below were carried out during the period January 1974 to December 1974:

1. An analysis of errors committed by respondents in the National Reading Performance Survey.

2. An inquiry into the reasons why certain errors were committed.

3. A study of the relationship of the reading tasks to decoding skills.

4. A study of the relationship of the reading tasks to a set of cloze tests.

5. A development and field test of an experimental test of basic reading competency in the schools.

6. An economic analysis of the data collected in the National Reading Performance Survey.

The Task Force on Essential Skills of the National Institute of Education was intimately involved in planning the tasks. Special acknowledgement is given to Monte Penney who has served as project monitor throughout the study, and to Marshall Smith and Donald Fisher.

In carrying out the six tasks listed above, many individuals at Educational Testing Service were involved. With the exception of the economic analysis task, responsibilities for the tasks were shared. Lola Rhea Appel, project manager, coordinated all activities and assisted in interviewing adults in the error inquiry, administering tests in the decoding and cloze tasks, and observing classrooms in the New York State experimental test administrations. The project director directed all phases of the study and is responsible for the over-all editing of the report.

In the error analysis phase of the study, Robert Patrick and Catherine Stasz assisted in the research which is reported in Chapter 1.

The task on decoding skills was directed by Carolyn Massad. She was assisted by Gertrude Conlan, Shirley Duby, Carol Ann Dwyer, Helen Fennimore, Jeanne Finelli, Dorothy Mazza, Gena Reisner, Frances Swineford, Spencer Swinton, Susan Valentine, and Janet Zahn, and is the author of Chapter 2 of this report.

In the task involving the relationship of functional reading to performance on cloze tests, Michael Zieky and Spencer Swinton assisted in designing the study and selecting the cloze tests. The project manager directed and carried out the testing of students in grades 7, 10, and 12 with the assistance of John Savarese, Spencer Swinton and Don Powers, and directed the scoring of the materials. Mr. Swinton and Mr. Zieky are the primary authors of Chapter 3 of this report.

An experimental test for students was developed in cooperation with the New York State Department of Education under the direction of the project director with the assistance of the project manager. The results of this task are reported in Chapter 4.

The economic analysis was designed and carried out by Educational Testing Service economists Kan-Hua Young and Dean Jamison. They are the authors of Chapter 5 of this report.

Marston Case, Kitty Kornhauser, Judy Pollock-Ohls, Robert Patrick, and Frances Swineford prepared the tapes and programs for analyzing the data in the various tasks. Eliza Kealy directed the typing and printing of this report with the assistance of Lucy Harvey, Claudia Modaffare, and Vicky Vinch.

The assistance of many persons outside Educational Testing Service is gratefully acknowledged:

In the New York State Department of Education, Jane Algozzine, Bureau of Reading; Mary Corrado, Division of Educational Testing; Thomas Fitzgerald, Bureau of Reading; Virginia Franseki, Bureau of Reading; Steven Kidder, Bureau of School and Cultural Research; Winsor Lott, Division of Educational Testing; Lois Mathison, Basic and Continuing Education; and Robert O'Reilly, Bureau of School and Cultural Research.

On the Advisory Committee for the decoding task, Dr. Ira E. Aaron, University of Georgia; Mrs. Jane Algozzine, New York State Department of Education; Dr. Don A. Brown, University of Northern Colorado; Dr. John B. Carroll, University of North Carolina; Dr. Jeanne Chall, Harvard University; Dr. Lynette Gaines, University of South Alabama; and Dr. Barbara Palmer, University of Connecticut.

In the Adult Learning Centers, Rose Lamorella and Garrett Murphy in Albany; Edwin Agresta and Kay Costopoulos in Schenectady; and Winona Peace, Judy Edalman, and Phyllis Spitzer in Yonkers.

In the school districts, James Crisafulli, English Department Chairman; Thomas J. Garberina, Curriculum Coordinator, Reading and Basic Instruction; M. Barbara Gilles, District Reading Director; John Grant, Assistant Superintendent; Doris Selub, Director of Reading; Terry Thompson, English Department Chairman; Hal Weller, Director of Reading; and Tim Xenos, Director of Reading.

Finally, the cooperation and enthusiasm of the many adults and students who agreed to participate in this study demand special recognition. The study could not have been completed without their assistance.

Richard T. Murphy
Project Director

CHAPTER 1

ERROR ANALYSIS AND INQUIRY

During the course of the Adult Functional Reading Study, two national surveys were conducted. In the first survey, a national probability sample (N = 5,096) of adults aged 16 and over was asked to describe the kinds of reading performed in the course of a normal day, the amount of time spent in carrying out the reading described, and the importance attached to the various kinds of reading performed. Results of this survey are published in Reading Activities of American Adults (Sharon, 1972) and Appendix A of Adult Functional Reading Study (Murphy, 1973). In the second survey, a national probability sample (N = 7,866) of adults aged 16 and over was asked to respond to 170 reading tasks representative of the kinds of reading reported in the reading activities survey. The 170 reading tasks were administered in 10 sets so that each adult responded to exactly 17 items. Results of the second survey are reported in Adult Functional Reading Study (Murphy, 1973).

In the National Reading Performance Survey, the survey administrators read the directions for each reading task orally to the respondents. The respondents then indicated their responses by underlining, circling, or placing an X on a portion of the stimulus material used in each reading task. The reading tasks were not administered in the common multiple choice format. Because of this type of administration, designed to correspond as closely as possible to a natural setting in which one might encounter such reading tasks, it was not possible to perform the ordinary type of error analysis with preconceived distracters listed as the only possible responses. In the materials used in the National Reading Performance Survey, a great variety of responses was possible. In order to analyze the kinds of errors made by respondents, a 10% random sample of respondents was chosen and their responses examined in considerable detail.

Erroneous Responses

In order to examine erroneous responses in detail, and to compare erroneous responses within tasks and across tasks, each response that had been coded as incorrect in the random sample was listed on a master copy of the stimulus material for each reading task. In addition, the number of adults giving a particular incorrect response was listed. The number of distinct incorrect responses ranged from as few as one incorrect response to as many as twenty-seven. The proportion of adults actually giving an incorrect response can be used as a measure of the "importance" or "strength" of that particular incorrect response.

To illustrate the kind of information that is available on the 170 reading tasks used in the National Reading Performance Survey, the results for five tasks are given on the following pages. The error percentages reported are the percentages of the total number of erroneous responses made on a particular item. Thus, they add up to 100%. It should be noted, however, that the numbers involved are small. In a 10% random sample, the number of respondents to any particular item is approximately 80. Therefore, the number giving incorrect responses is correspondingly small. For example, in the summary for Book 1: Item 13, the number of persons involved is only 19. This should be kept in mind in interpreting these data. The difficulty level from the national survey and the number of respondents in the 10% random sample are listed for each reading task.

Illustrative Examples

1. Book 4: Item 2 (Difficulty level = .999; Sample N = 77)

   a. Oral direction: Place a circle around the bottle of liquid that would be safe to drink.

   b. Errors: NONE

2. Book 1: Item 13 (Difficulty level = .666; Sample N = 76)

   a. Oral direction: Look at the train schedules. Put a circle around the time the daily train leaving Trenton at 1:46 P.M. arrives in Washington.

   b. Errors:

A - 42.1%	E - 15.8%
B - 5.3%	F - 10.5%
C - 5.3%	G - 5.3%
D - 5.3%	H - 5.3%
	I - 5.3%

3. Book 3: Item 1 (Difficulty level = .957; Sample N = 77)

   a. Oral direction: Put a circle around the label that would be the best one to put on a box used to mail something easily broken.

   b. Errors: A - 50%
      B - 50%

4. Book 4: Item 9 (Difficulty level = .814; Sample N = 77)

   a. Oral direction: Look at the garment tags. Circle the two tags that indicate the garments are made from 100% Polyester.

   b. Errors: A - 35.7% B - 7.1% C - 21.3% D - 7.1% E - 21.3% F - 7.1%

5. Book 5: Item 3 (Difficulty level = .928; Sample N = 77)

   a. Oral direction: Look at the application for employment. Put an X in the space where you would write the name and address of someone to notify in case of emergency.

   b. Errors: A - 40% B - 20% C - 20% D - 20%

Analysis of Errors

The number of distinct errors on the 170 reading tasks ranged from as few as one to as many as 27. Approximately 900 distinct errors have been identified and actually illustrated on a master set of reading tasks delivered to the National Institute of Education. Some errors were made by as few as one person; others were made by as many as 15 or 20 persons. At least one error was common to 55 persons. No general hypotheses were formulated in developing the materials. These are more or less complex reading tasks found in ordinary life. The empirical results give reasonable indications of the percentages of adults, and the percentages of approximately 40 subgroups of adults, who can answer correctly simple questions based on these materials. The error analysis is an attempt to go further -- an attempt (1) to discover the kinds of errors made by respondents who can not answer these questions correctly and (2) to determine why particular errors are made.

After the errors identified in the 10% random sample of respondents had all been listed on master copies of the reading tasks, an attempt was made to categorize the kinds of errors made. Several different classification schemes were tried. However, the great variety of errors that appears in this set of real responses seems to defy simple classification. It is possible to identify many errors that are questionable. That is to say, the open response format allows respondents complete freedom in indicating their


AMMONIA
MILK
FLOOR
WAX

New York—Washington

THE TIME from 12:01 am to 12 o'clock noon inclusive is indicated by light face type, from 12:01 pm to 12 o'clock midnight, inclusive, by dark-face type

	New York, N.Y. (Penna. Sta.) Leave	Newark, N.J. Leave	Trenton, N.J. Leave	North Philadelphia, Pa. Leave	Philadelphia (Penn Central Sta.—30th St.) Leave	Wilmington, Del. Leave	Baltimore, Md. Arrive	Capital Beltway, Md. Arrive	Washington, D.C. Arrive
177 Mondays thru Saturdays	3:23 AM	3 38	4.27	5 00	5.09	5 48	6 59	—	7:50 AM A
131 Mondays thru Saturdays	6 30 AM	6 45	7 32	8 00	8 10	8 38	9 39	—	10 20 AM
101 **Metroliner** Mondays thru Fridays	7 30 AM	c 7 42	—	—	8 43	—	9 49	d 10 13	10:25 AM
133 Daily	8:00 AM	8.16	8 59	9 26	9 35	10 18	11 18	—	12:00 Noon F
103 **Metroliner** Daily	8 30 AM	c 8.42	9 18	—	9:46	10 10	10 57	—	11 30 AM
135 Daily	9 30 AM	9 46	10 31	11 00	11 10	11 45	**12:48**	—	**1:30 PM**
137 Daily	10:45 AM	11 01	11 45	**12:12**	**12:21**	**12:51**	**1:51**	—	**2:40 PM**
105 **Metroliner** Daily	11 30 AM	c 11.42	**12:18**	—	**12:46**	**1:10**	**1:57**	—	**2:30 PM**
171 Daily C	**12:45 PM**	**1:01**	**1:46** E	**2:13**	**2:22**	**2:53** G	**3:58**	d **4:27**	**4:45 PM**
107 **Metroliner** Daily	**1:00 PM**	c **1:12**	—	—	**2:13**	**2:36**	**3:23**	d **3:47**	**4:00 PM**
163 Runs Feb 12 and 15 only	**2:00 PM**	**2:16**	**3:00**	**3:29**	**3:40**	**4:09**	**5:10**	—	**5:50 PM**
173 Daily	**3:00 PM**	**3.16**	**4:00**	**4:29**	**4:39**	**5:08**	**6:09**	—	**6:50 PM**
109 **Metroliner** Daily	**4:15 PM**	c **4.27**	**5:03**	—	**5:31**	**5:55**	**6:42**	—	**7:15 PM**
165 Runs Feb 12 and 15 only	**4:30 PM**	c **4:45**	d **5:28**	**5:55**	**6:15**	**6:43**	**7:44**	d **8:14**	**8:30 PM**
111 **Metroliner** Sundays thru Fridays	**5:00 PM**	—	—	—	**6:10**	**6:33**	**7:17**	d **7:40**	**7:55 PM**
175 Daily	**5:45 PM**	**6:01**	**6:48**	**7:15**	**7:24**	**7:52**	**8:53**	—	**9:35 PM** D
159 Sundays only	**6:30 PM**	**6:46**	**7:29**	**7:57**	**8:07**	**8:36**	**9:40**	d **10:09**	**10:25 PM**
139 Mondays thru Saturdays	**6:30 PM**	**6:45**	**7.38**	**8:07**	**8:26**	**8:54**	**10:00**	d **10:40**	**10:55 PM**
155 Daily	**7 30 PM**	**7:46**	**8:29**	**8:57**	**9:06**	**9:39**	**10:40**	—	**11:20 PM**
113 **Metroliner** Sundays thru Fridays	**8:30 PM**	c **8:42**	—	—	**9:43**	**10:06**	d **10:53**	d **11:16**	**11:30 PM**
147 Daily	**9:00 PM**	**9:15**	**10:04**	**10:41**	**11:01**	**11:29**	12 37	—	1 35 AM
161 Sundays and Feb 15 will not run Feb 14	**10:00 PM**	**10:16**	**11:05**	**11.33**	**11:46**	12 22	1 29	—	2 15 AM

H

Washington — New York

	Washington, D.C. Leave	Capital Beltway, Md. Leave	Baltimore, Md. Leave	Wilmington, Del. Leave	Philadelphia, Pa. (Penn Central Sta.—30th St.) Leave	North Philadelphia, Pa. Leave	Trenton, N.J. Leave B	Newark, N.J. Arrive	New York, N.Y. Arrive
140 Daily	2 25 AM	—	d 3 05	d 4 14	d 4 50	d 5 23	d 5 50	d 6 39	7 00 AM
170 Daily	6 55 AM	—	7 36	8 37	9 10	9 20	9 50	10 38	10 55 AM
100 **Metroliner** Mondays thru Fridays	7 30 AM	s 7 40	8 06	8 51	9 15	—	—	d 10 16	10 30 AM
102 **Metroliner** Daily	8 30 AM	—	9 02	9 47	10 13	—	10 39	d 11 16	11 30 AM
126 Daily	8 40 AM	c 8 54	9 25	10 29	10 58	11 07	11 35	**12:22**	**12:38 PM**
172 Daily	10 00 AM	—	10 40	11 41	**12:14**	**12:23**	**12:51**	**1:35**	**1:50 PM**
130 Daily	11 40 AM	—	12 21	**1:35**	**2:03**	**2:15**	**2:45**	**3:30**	**3:45 PM**
104 **Metroliner** Daily	12 00 Noon	s 12:11	**12:37**	—	**1.42**	—	**2:09**	d **2:46**	**3:00 PM**
106 **Metroliner** Daily	**1:00 PM**	—	**1:32**	**2:17**	**2:43**	—	**3:09**	d **3:46**	**4:00 PM**
174 Daily	**1:40 PM**	—	**2 21**	**3 22**	**4:00**	**4:10**	**4.39**	**5:24**	**5:40 PM**
132 Daily	**3:00 PM**	c **3:14**	**3:45**	**4:45**	**5:13**	**5.22**	**5:51**	**6:34**	**6:50 PM**
152 Daily	**4:00 PM**	c **4:14**	**4:44**	**5:52**	**6:19**	**6:28**	b **6:55**	**7.40**	**7:55 PM**
108 **Metroliner** Daily	**4:30 PM**	—	**5:02**	**5:47**	**6:13**	—	**6:39**	d **7:17**	**7:30 PM**
154 Sundays thru Fridays	**5:00 PM**	—	**5 47**	**6.50**	**7.19**	**7:29**	**7:56** I	**8.40**	**8:55 PM**
110 **Metroliner** Sundays thru Fridays	**6:00 PM**	—	**6.32**	**7:17**	**7:43**	—	**8:09**	d **8:46**	**9:00 PM**
166 Saturdays Sundays and Feb 15 will not run Feb 14	**6:05 PM**	—	**6 45**	**7 50**	**8 19**	**8:29**	**9:03**	**9 55**	**10:10 PM**
158 Daily	**7 25 PM**	c **7 38**	**8:17**	**9 12**	**9 42**	**9:58**	**10 26**	**11.20**	**11 35 PM**
112 **Metroliner** Sundays thru Fridays	**8:30 PM**	—	**9:02**	**9:47**	**10:13**	—	**10:39**	d **11:16**	**11:30 PM**
176 Daily	**10:15 PM**	—	**10.55**	12 04	12 51	1 01	1 31	2 24	2 49 AM

Reference Notes
b Stops Mondays thru Saturdays
c Stops only to receive passengers
d Stops only to discharge passengers
s Stops Mondays thru Fridays to receive passengers
★ Meals and Beverages served at seats

FRAGILE — HANDLE WITH CARE

THIS END UP

CONFIDENTIAL

PERISHABLE — KEEP REFRIGERATED

A

B

SPECIAL DELIVERY

AIR PARCEL POST

Book 4: Item 9

E

A

100% ARNEL® TRIACETATE

ARNEL® is a licensed trademark which can be used to identify Celanese Triacetate Fibers only in fabrics of proven performance. Celanese laboratories continually evaluate fabrics of the same style as in this garment against rigorous performance standards. If the fabric styles fail to meet these standards, the Triacetate fibers cannot be identified as ARNEL.

Machine wash - gentle cycle - warm water.
- Tumble dry, warm setting.
- Steam iron at medium setting.
- Wash dark colors or prints separately.
- Wash pleated garmets by hand - drip-dry - do not wring.

T0009-3

A

B

100% TEXTURED POLYESTER
- Machine Wash and Dry-Wash and wear cycle
- Will Not Shrink or Stretch
- Wrinkle Resistant
- NEEDS NO IRONING
- DO NOT DRY CLEAN

PALISADE KNITTING MILLS, GUTTENBERG, N.J.

F

C

6027 SIZE 9
65% ACETATE
35% FORTREL
Polyester

MACHINE WASH
PRESS ON
WRONG SIDE
WITH COOL IRON
~~P O 111951~~
WPL 82311

D

D

WASHING INSTRUCTIONS

-Wash by machine at low temperature. Dry by machine at low temperature or hang on line. If ironing is desired, use cool iron only.

ADVANTAGES.
1. Smooth comfort: ideal for outerwear or next to skin.
2. Excellent dimensional stability.
3. Delightful warmth with light weight.
4. Easy-care, excellent washability.
5. Mothproof and mildewproof.

The irregularities and variations in the weave of this fabric are characteristics of this fabric and are in no way to be considered as defects.

B

E

RN 26559
MADE IN U.S.A.
06
100% POLYESTER C
MACHINE-WASHABLE
DRY-CLEANABLE

Lot 5218A Style 3551P
10

SUGGESTED RETAIL PRICE
$13.00

Book 5: Item 3

APPLICATION FOR EMPLOYMENT DATE

[redacted] (LAST) (FIRST) (MIDDLE) (MAIDEN)

NAME

LOCAL ADDRESS X (B) TELEPHONE NO.

ZIP CODE

HEIGHT WEIGHT

PERMANENT ADDRESS X (C)

MARITAL STATUS SINGLE ☐ MARRIED ☐ DIVORCED ☐ WIDOWED ☐ SEPARATED ☐

NAME & ADDRESS OF PERSON TO NOTIFY IN EMERGENCY | NO. OF CHILDREN | AGES:

FIRST NAME OF SPOUSE X (A) PLACE OF EMPLOYMENT

EDUCATION	NAME & LOCATION OF SCHOOL	FROM	TO	COURSE OR MAJOR	YEAR GRAD	DEGREE
HIGH SCHOOL	X (D)					
COLLEGE						
BUSINESS OTHER SCHOOLS						
SPECIAL STUDY IF ANY						

responses. As a result, there is a sizable group of errors that can be usefully classified as inclusive errors. The portion of the stimulus material underlined or circled does in fact contain the correct information, but so much additional information is included that the response was judged to be incorrect. In cases where the portion circled or underlined included an alternative response that was clearly incorrect, it seems clear to consider the response as incorrect. In cases where the additional information circled or underlined is not relevant to the correct response, it is more difficult to determine that the respondent did not know the correct answer to the question. He simply may not have followed the directions carefully. Thus, there is some ambiguity in this type of response and the judgment of error is open to question. This is a considerable problem in working with these real life tasks with open end responses. However, to keep a proper perspective, this kind of error accounts for less than 10% of the total number of errors.

A second class of errors can be identified, though the decision that a particular erroneous response belongs in this category is somewhat arbitrary. These are responses that are related to, but do not include, what the item developer considered to be the correct response. For example, a question asks the respondent to circle the telephone number that should be called under certain circumstances. Instead of circling the telephone number, a respondent circles the person or place which is uniquely connected with the correct telephone number. Although the directions were not followed explicitly, the respondent probably does understand the question and probably would dial the correct telephone number in a simulation of this activity. However, to once again attain a proper perspective, this kind of response, based on the judgments of the project staff, occurred in only about 5% of the erroneous responses.

By and large, more than 85% of the erroneous responses appear to be clearly and unambiguously incorrect. Without actually interviewing the respondents to ascertain what they were thinking, it is difficult to categorize these responses. However, in an attempt to understand why adults give particular erroneous responses on the real life reading tasks, we asked a group of adults to volunteer to respond to the materials and then discuss their responses with two members of the project staff.

<u>Error Inquiry</u>

Two adult learning centers in New York State agreed to participate in this phase of the study. Approximately 100 adults were asked to respond to the materials which were group administered to small classes by the project staff. All of the students who participated volunteered to do so and were all adults over 16, male and female, and white, black, Spanish, or Oriental. The same 10 test booklets of 17 items each were used in the Learning Centers that had been administered in the National Reading Performance Survey. However, the question for each reading task appeared above the stimulus and was read silently by the respondents. There was no time limit to the test administration and most of the students completed the materials within 30 minutes. The reading item booklets were scored immediately and each respondent was interviewed individually by a project staff member as soon after taking the test as possible. In addition to an overall discussion of the materials in general, the students were asked to elaborate on their answers to the reading items. They were asked to explain why they answered as they did, if possible. The length of the interviews varied. In some cases, the interview lasted more than an hour because of the respondent's interest and involvement in the reading items.

Two very simple causes of incorrect responses emerged. Many persons simply do not understand the words in the reading tasks, and, when they do understand the individual words, they do not understand how to extract information from the forms in which these words are patterned. Patterns do not refer to the patterns of sentences and paragraphs, but rather to the patterns of forms, tables, and charts that are so commonly used in such everyday things as bills and schedules. This is a parsimonious explanation of lack of success in handling common reading tasks. Many poor readers are simply confused by such words and phrases as "to call up," "transplant," "apparel," "commencement," "lives," "toll," "injection," "series," "creed," "misstatement of fact," "confronts," "recipe," "ingredients," "lever," "firearms," "locker," "escape," "extinguisher," "ingestion," "correspondent," "to fill in," "came together," "permanent," "pesticide," "circle," "fourth," "to operate," "classified," "fuel," "stance," "minimum," "severe," "mild," "whom," and "experience." These are words included in the everyday reading tasks used in the study.

In addition, many persons who do have a sufficient vocabulary to understand the words included in the reading tasks do not understand how to handle simple words, phrases, and numbers when they are presented in rows and columns as in doctor's bills, telephone bills, train schedules, guides, report cards, application forms, election ballots, employment benefit forms, income tax forms, social security forms, and traffic tickets. Many respondents are simply confused by such materials, and, in many instances, responses are simply guesses and not the result of an intricate process which leads to an explainable response as an alternate to a correct one.

In some few instances, a particular incorrect response is common to a sizable group of persons and a reasonable explanation can be given for the response. For example, in the illustrative example, Book 1: Item 13, note that the incorrect response lettered A was given by about half the respondents who responded incorrectly. In the error inquiry, respondents understood the words "arrive" and "Washington" but did not understand how to use the train schedule. The 7:50 A.M. response was the response just below those two words on the page.

Many responses do not appear to be amenable to explanation. In fact, in the error inquiry we discovered that reasons for particular responses to the reading tasks could be quite unique to the individual respondent. A few of these are given below.

1. In a list in which the respondent was to choose an entry corresponding to baby's clothes, the entry hampers appeared. A respondent who chose that entry explained that he thought hampers might be like "Pampers" -- a commercial product of disposable baby's diapers.

2. A list contained several amounts of alcohol and the effects associated with drinking such amounts.. A respondent was asked to circle the amount associated with a given effect. He circled a greater amount and gave as the reason his disagreement with the chart. He judged that the effect would be associated with a greater amount of alcohol.

3. A doctor's bill listed the amount owed. A respondent circled a higher amount listed elsewhere on the bill because it corresponded more closely to her own latest doctor's bill.

Discussion of Results

There appear to be two ways of considering these results. The vocabulary and formats employed in everyday reading materials are fixed and, therefore, children and young adults should learn, perhaps while still in school, to understand such vocabulary and formats. There is evidence that educators are moving in this direction in attempting to alert schools to these problems and to develop materials for use in the schools for teaching this vocabulary and these skills. However, the information may also be considered in reference to the common, everyday reading materials themselves. In developing such materials, producers may be encouraged to use simpler vocabulary and simpler formats. A large and complex table can certainly summarize a great deal of information. However, if such tables are difficult to understand for sizable groups of people, the saving in space may be of questionable benefit. Voting forms, tax forms, doctor's bills, guides, and schedules of all types can perhaps be simplified by expanding the materials if the benefit in comprehension is evident.

A final point should be made about the actual directions employed in administering the reading tasks in the national survey. The words and phrases identified in the reading tasks as difficult for some adults were also used in the directions. Since the intent of the question was to find out if the respondent could read the stimulus material, the questions should have been phrased in the simplest language. As a result of the error analysis and the error inquiry, a number of suggestions for re-phrasing the oral directions can be made for any possible further use of these materials. A set of alternative directions for some of the reading tasks has been deposited with NIE. Finally, the ambiguous nature of some of the reading tasks has become evident from the results of the error analysis and inquiry. A list of such tasks is also on file at NIE.

CHAPTER 2

RELATIONSHIP OF DECODING TO ADULT FUNCTIONAL READING

In this task, a subset of the adult functional reading items from the National Reading Survey was administered to a group of approximately eighty adults together with a set of materials designed to assess decoding skills. The purpose of the task is to examine whether or not the inability to answer correctly the adult reading tasks is attributable to inadequate decoding skills.

PREPARATION AND ORGANIZATION FOR THE STUDY

The Advisory Committee

Since the term "decoding" is conceptualized differently by various people, the first step undertaken by the investigators was to convene an advisory committee of specialists in reading with the objective of defining "decoding" for the purposes of this study. Further, it was expected that the advisory committee would assist in determining how decoding should be tested in adult subjects.

The meeting was held June 29 - 30, 1974 at Educational Testing Service (ETS) in Princeton, New Jersey. The advisory committee consisted of:

Ira E. Aaron, University of Georgia
Jane Algozzine, New York State Department of Education
Don A. Brown, University of Northern Colorado
Jeanne Chall, Harvard University
Lynette Gaines, University of South Alabama
Barbara Palmer, University of Connecticut

John B. Carroll, then of ETS and currently at the University of North Carolina, also attended the meeting in a consultative role. Committee members were selected from a substantial list of candidates and selection was based on the aim of obtaining committee members who were familiar with and understood the problems of adult literacy. Further, every attempt was made to have the committee membership reflect as much as possible the wide range of viewpoints that exist in the field of reading.

First, the committee as a whole objected to the loose use of the term "decoding" and felt that, for purposes of the task at hand, <u>word attack skills</u> was a more appropriate concept from and with which to work. The committee then focused on identifying the different elements involved in decoding and specifically those elements related to assessing the ability of adults to decode. Five "guide questions" were proposed as an aid to the committee's discussion:

1.) Is decoding a problem in the performance of adults on the 170 tasks in the National Reading Performance Survey?
2.) If so, what elements present major decoding problems?
3.) What are the characteristics of these elements and can they be classified?
4.) Can these be ordered sequentially and, if so, on what basis?
5.) Is there a test or can one be generated to determine where adults fall out on this(these) sequence(s)?

The committee considered the possibility of using existing tests and test specifications appropriate for the adult level either as the actual instruments for providing information or as guidelines to aid in the development of a new test on word attack skills. Some materials and test copies had been gathered from the ETS Test Collection for the purpose of providing the committee with resources should they wish to review materials. In discussion, the committee felt that the existing decoding tests, while partially adequate, could not meet all of the needs of the present study.

It was decided that, while the committee could borrow ideas from existing tests, a complete list of specifications should be drawn up and a new test designed to incorporate these. The initial list of areas to be considered included: visual discrimination; auditory discrimination; basic sight vocabulary; phoneme-grapheme correspondence for initial consonants, ending consonants, short vowels, long vowels, diphthongs, and digraphs; context clues; structural analysis; syllabication; blending. The formulation of this list led the committee to discuss both how the list could be used in the development of specifications and how the areas listed could be used to provide valid information.

The committee felt that prerequisite to any testing for determining adult literacy would be to find out if the adult could identify letters of the alphabet. This ability could be tested by having the adult point out to the examiner specified letters, match letters to sounds said by the examiner, and so on. Recognition of both upper and lower case letters in printed form should be required.

The committee also felt that, to further determine whether adults could decode, a series of graded paragraphs should be used. These paragraphs would be used in a one-to-one testing situation and the adult would read them orally at sight. The adults could read the various paragraphs until errors were so excessive as to preclude any comprehension of content, i. e., until the adults reached their frustration level. After this point, the examiner might just ask if the adult could read any of the words from the higher level paragraph(s).

Each of the "hierarchical" paragraphs to be used in the decoding test would have accompanying exercises based on the word attack skills needed to read that passage. The committee listed word attack skills essential to mastery of decoding in a hierarchical form for this purpose. Specific instructions as to which elements would be tested along with each paragraph and how they should be tested were provided. However, the committee felt that successful oral reading of the paragraphs would indicate that the adult has the ability to decode. In this way, it was felt that concurrent validity would be established.

The sequence of oral sight-reading materials was to consist of a minimum of three paragraphs, the first of which would approximate the upper first-grade level and have accompanying questions checking the lowest level of word attack skills, such as visual and auditory discrimination. The second paragraph would approximate the upper second-grade or early third-grade level and the accompanying skills questions would cover such things as initial and ending consonant sounds-symbols, long and short vowel sounds-symbols, and diphthongs. The third paragraph would approximate middle or upper fourth-grade difficulty and the accompanying skills questions would cover such things as blending, syllabication, and some upper level structural analysis skills. Such things as basic sight vocabulary and use of context clues would cut across all levels of paragraphs.

The committee indicated that if an adult could successfully read at sight the three graded paragraphs and do most of the word attack skills incorporated in them, he could be said to have no major decoding problems. The adult should then be given what would be considered a fifth- or sixth-grade level passage to sight read orally. This passage would be taken from the 170 tasks developed for the National Reading Survey.

Suggestions for scoring the oral reading of the various paragraphs were also provided. The final specifications for the decoding instrument as well as the test administration procedures are contained in the Appendix for this chapter.

If an adult should still experience problems with reading after scores on the decoding tests indicated a mastery of decoding skills, it was suggested that an appropriate standardized diagnostic reading test be given to determine what other kinds of difficulties the adult might have that made him unable to read adequately.

Before the advisory committee meeting adjourned, the committee members made it clear that the proposed instrument for this study was to be a research instrument. It was felt that an instrument similar to the one proposed was needed in the field but that a thorough investigation of the entire area of assessing decoding in adults was needed before a final instrument should be developed. The proposed decoding measure would be only a first step in the direction of meeting an existing need.

Instrument Development and Preparation

Following the advisory committee meeting, the development of the decoding instrument was undertaken. All test questions, paragraphs to be read orally, and test administration procedures were written and reviewed by ETS staff with expertise in the field of reading assessment. Materials and suggestions contributed by several committee members provided the bases for the materials written. Each item and paragraph, along with the administration procedures for the same, were reviewed by three staff members other than the writer. Further, the ETS editors reviewed the material for clarity of expression, to eliminate any existing ambiguities, and so on.

In addition to the decoding instrument, 33 functional reading tasks were selected from the 170 administered in the National Reading Survey. These 33 tasks were prepared in the form of a reading test, the results of which would be compared to the results of achievement on the decoding instrument.

DESCRIPTION OF THE INSTRUMENTS USED

The Survey of Reading Materials

Instrument Description. A subset of the 170 reading tasks from the National Reading Survey was chosen and fashioned into a test estimated to require about an hour of respondent time. An attempt was made to select the least ambiguous items on the basis of the error analysis and error inquiry. In addition, items that employed in the directions exact words appearing in the stimulus materials were not used. The rationale here was to prevent simple matching of words and to further justify using printed directions to facilitate group administration of the materials. Furthermore, an attempt was made to approximate the distribution of item difficulties and kinds of forms and benefits present in the national survey. The distribution of items by difficulty and descriptors is given in Table 2.1.

Table 2.1

DESCRIPTIONS OF READING TASKS

Difficulty Level	# of Items	Form	# of Items	Benefit	# of Items
90-100	13	Book	1	Economic	6
80-89	8	Periodical	2	Occupational	2
70-79	7	Legal Document	3	Education/Culture	3
60-69	2	Listing	5	Recreation	3
50-59	1	Instruction	6	Health	6
40-49	2	Advertising	4	Maintenance	3
		Form	2	Personal Relationships	4
		Personal Communication	5	Citizenship	6
		Miscellaneous	5		

In all, 33 reading tasks were incorporated into an experimental test of functional reading. As in the national survey, respondents had to indicate their answers by underlining, circling, or placing an X on a portion of the stimulus material. There was no fixed time limit; each respondent was to be able to proceed at his own rate. The test could be either individually or group administered.

Scoring. The reading tasks were scored right, wrong, or omit using the coding specifications developed for coding the national survey materials. The coding was performed by the director of the coding phase of the national survey. The project director checked 20% of the materials for quality control. The scores were then coded on a machine scorable answer sheet for analysis.

The Decoding Instrument

Instrument Description. The decoding instrument developed reflects the specifications provided by the advisory committee. Three tests, two of which have been divided into sections, made up the decoding instrument. Table 2.2 provides a summarized description of the decoding instrument.

Table 2.2

DESCRIPTION OF THE DECODING INSTRUMENT

Test	Number of Items*	Item Type
Alphabet Test	26	4- to 8-choice multiple-choice
Test of Word Attack Skills	66	(See Below)
Level A	13	2- to 6-choice multiple-choice
Level B	32	3- and 4-choice multiple-choice
Level C	21	3- and 4-choice multiple-choice plus one 1-word written response and one oral reading of three wrods**
Test of Oral Reading	4	Passage reading orally

Note: *The total score on each test equaled the number of items correct.

**The oral reading of two words was administered with the Test of Oral Reading due to the nature of the item.

For the Alphabet Test, each subject had a test booklet in which to respond. Each of the 26 items consisted of a row of four to eight letters. A mixture of upper- and lower-case letters was used for each item. The test administrator said one of the letters in the row and the subject was to put an X on that letter. The letter was said only once unless conditions were such that it was difficult for the subjects to hear what was said and, then, the letter was repeated. There was no fixed time limit.

Each subject had a test booklet in which to respond to items in the Test of Word Attack Skills, the exception being the one item that was administered with the Test of Oral Reading. The test had no fixed time limit. To the extent possible, every subject was given sufficient time to attempt every item.

The Test of Word Attack Skills had three sections, each one representing one of the three levels (A - C) specified by the advisory committee. Level A consisted of 13 items divided into four sets, each of which was to measure a specified area and was preceded by an example. The first, second and fourth set of items each had three items and were intended to measure auditory discrimination, visual discrimination and the use of context clues respectively. The third set contained four items and was intended to measure recognition of basic sight vocabulary. The Administrator's Manual specified required procedures for each item.

Level B consisted of 32 items divided into five sets each of which was preceded by one or two examples. The first set contained five items: four to measure recognition of initial consonant sounds-symbols; one for measuring recognition of digraphs occurring in the initial position. The second set contained five items: four to measure recognition of final consonant sounds-symbols; one to measure recognition of digraphs occuring in the final position. The third set containing 12 items was to measure recognition of short and long vowel sounds-symbols as well as recognition of diphthong and digraph sounds-symbols. The fourth set, with four items, was intended to measure the use of context clues. The fifth set had six items to measure the use of structural analysis. For all but the fourth set of items, the Administrator's Manual specified required procedures for each item. The directions were given, the example(s) done and the subjects proceeded to work on their own for the fourth set of items.

Level C consisted of six sets of items, five of which contained a total of 20 items and were in the test booklet labeled Test of Word Attack Skills. The last set contained only one item and was administered with the Test of Oral Reading due to the nature of the item. Each set of items except the first was preceded by an example. The first set contained five items and was similar to the last set in Level B in that the purpose of the set was to measure the use of structural analysis; however, the Level C items were intended to be more difficult. No example was needed since similar items had preceded this set. The second set contained five items for measuring recognition of digraphs. The third set, consisting of four items, was intended to measure the use of context clues. The fourth and fifth sets, each containing three items, were to measure syllabication and blending syllables, respectively. For the first and fifth sets of items the Administrator's Manual specified required procedures for each item, but directions were given, the examples done and the subjects proceeded to work on their own for the remaining sets of items. For the Test of Word Attack Skills, Table 2.3 indicates for each level: the specification area tested, the number of items per area, and the item numbers as they occur in the test booklet.

Table 2.3

DESCRIPTION OF THE TEST OF WORD ATTACK SKILLS

Level	Specification Area	Number of Items Per Area	Item Numbers
A	Auditory discrimination	3	1-3
	Visual discrimination	3	4-6
	Recognition of basic sight words	4	7-10
	Use of context clues	3	11-13
B	Recognition of initial consonant sounds	4	14-17
	Recognition of final consonant sounds	4	19-22
	Recognition of short vowel sounds	5	24,26,28 31,33*
	Recognition of long vowel sounds	5	25,27,29* 30*,35*
	Recognition of diphthongs	2	32,34
	Recognition of digraphs - both vowel and consonant	6	18,23,29 30,33,35
	Use of context clues	4	36-39
	Use of structural analysis	6	40-45
C	Use of structural analysis	5	46-50
	Recognition of digraphs - all	5	51-55
	Use of context clues	4	56-59
	Syllabication	3	60-62
	Blending syllables	3	63-65

*Items indicated overlap with recognition of digraphs.

The booklet labeled Test of Oral Reading contained three parts. The first part consisted of one item for which the subject was asked to orally read, at sight, three words that were divided into syllables. This item was to measure blending syllables and was classified as the sixth set of Level C of the Test of Word Attack Skills. The remainder of the booklet was the Test of Oral Reading that included four passages to be read aloud, at sight. First, there was a part containing three passages, each one representing one of the three levels (A - C) and reflecting the skills specified for that level in the hierarchy indicated by the advisory committee. The three passages were followed by a fourth passage that was taken from one of the 170 tasks in the National Reading Performance Survey. The four passages in the Test of Oral Reading were intended to be of increasing difficulty from the first through the fourth. The readability of each passage was determined by using two formulas (Dale and Chall, 1948; Gray and Leary, 1935). The results showed that each passage written and the one from the National Reading Survey approximated the level they were intended to be per the advisory committee's recommendations.

All three tests in the decoding instrument are appropriate for individual administration. Both the Alphabet Test and the Test of Word Attack Skills can be group administered but the Test of Oral Reading must be individually administered.

Scoring. Each item was scored on a pass or fail basis. That is, if a subject's response was correct, he received a "pass" on the item. If his response was incorrect or he made no response, the subject received a "fail" on the item. Every subject had an opportunity to try every item in all three tests of the decoding instrument.

Using test keys indicating the one correct answer, four scorers scored the subjects' Alphabet Test and the Test of Word Attack Skills. For quality control of scoring, one-half of the test booklets, randomly selected, were scored by two independent scorers and there appeared to be no problems in scoring. The score for each item was coded on a machine scorable answer sheet.

Scoring for the oral items was done when the items were administered. There were four test administrators who gave and simultaneously scored the items. Prior to the testing sessions, the test administrators were trained and discussed any questions regarding the administration and/or the scoring procedures. Due to the nature of the items it was difficult to determine interscorer reliabilities; however, every indication was that the scoring was consistent across scorers. Standard guidelines for scoring the oral items were in the Administrator's Manual and followed by each administrator.

The item classified with the Test of Word Attack Skills and intended to measure blending syllables was to be scored using the following guidelines:

No errors are allowed. All three words must be read correctly.

ERRORS include:

Omitting a part of the word.

Replacing correct "sounds-symbols" with incorrect "sounds-symbols".

Adding "sounds-symbols" not present in the words.

Changing the position of "sounds-symbols" to incorrect positions in the words.

Mispronouncing the word. In the words "innocent" and "newstand", the last letter may be dropped due to dialect. Do not count this as an error.

NOT ERRORS are such things as:

Repeating the word or parts as in stuttering.

Pausing before reading the word or syllable.

The first three passages in the Test of Oral Reading were scored as follows:

Scoring Criteria

Passage for Level A - one error only is allowed

Passage for Level B - one error only is allowed

Passage for Level C - two errors only are allowed

ERRORS include:

Omitting a word except a word or article (e.g., a, the) that appears more than once.

Substituting another word for the one there.

Adding a word or words that are not there.

Changing the position of words (e.g., is it for it is).

Mispronouncing sight words or words that follow regular rules of pronunciation. Words not following rules are in passage B: "listen"
in passage C: "weather"

Asking for aid in pronouncing a word that the subject tried and could not get.

NOT ERRORS are such things as:

Repeating words read.

Pausing briefly between words or sentences.

Ignoring punctuation marks in reading orally.

The scoring of the fourth passage involved the following guidelines:

Scoring Criteria

Only two errors are allowed.

ERRORS include:

Omitting a word except a word or article (e.g., a, the) that appears more than once.

Substituting another word for the one there.

Adding a word or words not there.

Changing the position of words (e.g., is it for it is).

Mispronouncing words unless corrected immediately following mispronunciation.

Asking for aid in pronouncing a word that the subject tried and could not get.

NOT ERRORS are such things as:

Repeating words read.

Pausing briefly between words or sentences.

Ignoring punctuation marks in reading orally.

TEST ADMINISTRATION

The Sample

The sample consisted of 77 adults enrolled in courses at an adult learning center. The learning center is housed in an old elementary school building, which is very well kept. There is a reading laboratory which has a specialist teacher supervising classes as the students use the facilities.

In order to get students to react freely and without stress during the testing sessions, it was decided not to obtain any personal information on the students. In this way, they would not feel threatened by thinking that their personal test results or information would be kept on record and used by others without their permission. However, it was noted that approximately one-third of the subjects were male, the female students being in the majority. In addition, the guidance counselors indicated that the classes were considered to be in self-contained classrooms except for the time spent in the reading laboratory. That is, except for the time spent in the reading laboratory, the subjects generally had only one teacher for all of their studies unless they attended classes for a whole day. If an adult attended classes all day, the teacher in the afternoon session was not the same as in the morning session since each teacher was assigned to teach for only a half-day. The subjects were drawn, in approximately the percentages indicated, from the following types of classes (as categorized by the guidance counselors):

- 8% Advanced Level of English as a Second Language
- 34% Students Achieving in Reading at the Primary-Grade Level
- 28% Students Achieving in Reading at the Intermediate-Grade Level
- 28% Students Achieving in Reading at the Pre-G. E. D. (High School Equivalency) Level

Given these classifications, it was hoped that the sample would range from very low-level readers to readers who would be considered quite literate. In other words, in order to satisfy the purpose of this study, an effort was made to obtain a sample with a wide range of decoding ability, including subjects with both major decoding problems and those who would be considered as having no major decoding problems.

Procedures in Administration

For purposes of this study, the Survey of Reading Materials, the Alphabet Test, and the Test of Word Attack Skills were group administered. The Test of Oral Reading was individually administered. The group administered tests were given at two sessions, each on a separate day, as follows:

Session 1 (Day 1)	Alphabet Test Test of Word Attack Skills
Session 2 (Day 2)	Survey of Reading Materials

For those absent on the days their class took the tests above, two other days for make-up of missed tests were scheduled. These make-up sessions were small group administrations with a few individual administrations where necessary due to student schedules.

Individual administration of the Test of Oral Reading covered a period of two days for four test administrators.

Before all testing sessions, both for the group and individually administered tests, the subjects were told that they should not become discouraged by any test questions. To avoid frustrating them, the subjects were told that not everyone was expected to be able to do all the questions although everyone was to have a chance to try to do every question in the test booklets. Also, it was made clear that in no way would the tests be used for grading them or would the test results go on their records. No time limits were set for any tests. Each subject proceeded at his own rate. As might be expected, some classes/individuals needed more time to go through the test booklets than did other classes/individuals.

RESULTS

Statistical Properties of the Instruments

The Survey of Reading Materials. The difficulty levels on the 33 reading tasks ranged from .33 to .97. The mean difficulty level for the entire test was .68. The mean difficulty level for this set of items in the national survey was .81. However, the percentage of non-responses on most of the items is considerably higher than the percentages in the national survey. In this sample, the percentages of non-responses ranges from 1% on item 1 to 48% on item 33. The average percent is 20%. In the national survey, the percent of non-responses on these items ranged from 0 to 17%. The average percent omitted was only 4%. The difference in non-response between the national sample and the sample for this study may be due to the fact that the directions, which were read orally to the respondents in the national survey, were not read orally to the respondents in this study. In the national survey it may have been difficult for a respondent not to attempt some response in the presence of the survey administrator.

For this study, the adults were given as much time as desired to complete the materials. The percent passing an item is based on the entire sample with no adjustments for omits. An omitted response is therefore interpreted as meaning that the respondent was not able to respond correctly. In some instances, he or she responded incorrectly. In other instances, he or she simply did not respond at all. However, for completeness, the percent of non-responses is given in Table 2.4. along with the percent passing and the r-biserials of the items. The scores on the total test ranged from a low of 1 to a high of 33. The mean score was 19.9 (S.D.=8.7). The reliability of the test using the Kuder-Richardson Formula 20 (Kuder-Richardson, 1937) is .94.

Table 2.4

ITEM ANALYSIS FOR THE SURVEY OF READING MATERIALS

Item No.	% Correct	% Non-response	r-biserials
1	97	1	*
2	90	5	0.65
3	97	4	*
4	86	7	0.56
5	89	5	0.74
6	85	4	0.73
7	77	8	0.74
8	86	9	0.63
9	59	4	0.69
10	63	15	0.58
11	89	7	0.53
12	60	4	0.68
13	90	7	0.70
14	56	17	0.61
15	81	8	0.45
16	70	9	0.37
17	59	12	0.47
18	70	15	0.57
19	70	19	0.60
20	57	31	0.68
21	59	23	0.47
22	79	17	0.37
23	50	31	0.68
24	79	28	0.58
25	51	37	0.62
26	65	36	0.66
27	61	37	0.71
28	54	39	0.65
29	33	36	0.63
30	41	33	0.55
31	48	45	0.70
32	54	47	0.64
33	49	48	0.38

*The r-biserial was greater than 1 due to lack of normality in the dichotomized variable and the inexact nature of the estimation procedure.

The Alphabet Test. This test proved to be a very easy test for this sample, 80% getting all 26 items correct. Another 12% had a total score of 25. The lowest score was 22, obtained by only two people in the sample.

Reliability of the Alphabet Test was determined to be .53 using the Kuder-Richardson Formula 20 (Kuder-Richardson, 1937) based on the internal consistency of all the items in the test. However, this figure does not adequately explain the data since between 95 and 100 percent of the sample passed all but three items, and more than 90% of the sample passed these three. The lack of variability in performance of the sample led to the low reliability calculated; this is not to say that the test was not a good measure of knowledge of the alphabet. In other words, little information could be obtained from the results of this test other than that almost the entire sample had mastered the alphabet.

The Test of Word Attack Skills. Table 2.5 gives the reliabilities for each level as well as for the total test, the Kuder-Richardson Formula 20 being used to determine the reliability coefficients.

Table 2.5

RELIABILITIES FOR THE TEST OF WORD ATTACK SKILLS

Test/Level	Reliability
Total Test	.91
Level A	.61
Level B	.86
Level C	.79

The lowest reliability coefficient was for Level A, all others indicating that the test as a whole and each of the other parts provided a reasonable measure of what was intended to be tested. Due to the lack of variability in the data for Level A, this part being very easy for the sample--as was the case for the Alphabet Test--and the fact that there were only 13 items, the reliability coefficient calculated was low. For this part of the test, seven items were passed by more than 95% of the sample. Only six items were sufficiently difficult in order to determine whether they contributed toward discriminating between high and low achievers. By examining the item analysis (Table 2.6) for Level A along with the specifications for the items (Table 2.3), these six items can be identified as the three items to measure auditory discrimination and three of the four items for measuring recognition of basic sight vocabulary, item 9 being the one item that was exceptionally easy.

Table 2.6

ITEM ANALYSIS FOR THE TEST OF WORD ATTACK SKILLS

Level	Item No.	Percent Passing	r-biserial for Part	r-biserial for Total Test
A	1	91	.66	.62
	2	91	.84	.60
	3	84	.85	.49
	4	99	*	*
	5	100	*	*
	6	99	*	*
	7	92	.76	.67
	8	95	.81	.68
	9	97	*	*
	10	95	.84	.75
	11	100	*	*
	12	97	*	*
	13	99	*	*
B	14	99	*	*
	15	97	*	*
	16	97	*	*
	17	97	*	*
	18	93	.62	.57
	19	88	.30	.37
	20	97	*	*
	21	99	*	*
	22	96	*	*
	23	91	.72	.78
	24	71	.65	.51
	25	91	.84	.78
	26	69	.55	.43
	27	31	.64	.56
	28	63	.61	.56
	29	76	.75	.74
	30	85	.60	.60

Table 2.6 (con't)

Level	Item No.	Percent Passing	r-biserial for Part	r-biserial for Total Test
B	31	36	.72	.73
	32	75	.76	.73
	33	17	.26	.09
	34	33	.69	.66
	35	60	.68	.61
	36	95	.78	.68
	37	96	*	*
	38	76	.31	.24
	39	97	*	*
	40	85	.56	.61
	41	64	.70	.60
	42	53	.64	.58
	43	80	.64	.63
	44	71	.82	.83
	45	67	.59	.59
C	46	48	.33	.26
	47	60	.48	.49
	48	81	.68	.70
	49	63	.38	.44
	50	80	.80	.87
	51	21	.40	.42
	52	29	.58	.55
	53	53	.52	.37
	54	32	.42	.35
	55	60	.70	.58
	56	95	.94	.94
	57	96	*	*
	58	89	.65	.64
	59	72	.48	.50
	60	81	.67	.59
	61	36	.59	.48
	62	69	.74	.67
	63	65	.75	.59
	64	57	.54	.46
	65	55	.63	.52
	66	77	.45	.46

*The r-biserial was greater than 1 due to lack of normality in the dichotomized variable and the inexact nature of the estimation procedure.

With respect to Level B, the item analysis (Table 2.6) shows that for this sample, the items intended to measure recognition of initial and final consonants were exceptionally easy for the sample and, therefore, did not contribute toward the discriminating power of this part of the test nor the total test. One exception to this was item 19 in which the subjects were asked to identify the grapheme r for the final phoneme in the word "water." It is to be noted, however, that the r-biserials for this item were .30 and .37, for Level B and the total test respectively, which shows that the discriminatory power was low and particularly so when compared to that of other items in the test. One other set of items in Level B, those for measuring the use of context clues, appeared to be very easy for this sample, one item being the exception. However, this item (38) did not appear to be discriminating very well between high and low achievers on the total test, although it may have been functioning at what might be considered a satisfactory level for the part of the test to which it belongs. Although another item (36) in this set of four items on the use of context clues was passed by 95% of the sample, it still seemed to have discriminating power for both Level B and the total test.

Level C had a lower reliability coefficient than Level B but this was most likely a function of the fact that Level C has 11 items less than Level B. In Level C, only one item (57) in the set of four items for measuring the use of context clues was exceptionally easy for the sample. Also, one item (46) in the set of five items intended to measure structural analysis did not seem to have a high level of discriminating power for Level C nor did it reach a satisfactory level of discriminating power for the total test.

The means and standard deviations for the parts and total Test of Word Attack Skills are given in Table 2.7. This table also shows the range of obtained scores in the sample as well as the total possible score.

Table 2.7

MEASURES OF CENTRAL TENDENCY AND DISPERSION OF SCORES ON THE TEST OF WORD ATTACK SKILLS

Test/Level	Total Possible Score	Range of Obtained Scores	Mean	Standard Deviation
Total Test	66	15 - 66	50.0	9.3
Level A	13	7 - 13	12.4	1.1
Level B	32	5 - 32	24.5	5.0
Level C	21	1 - 21	13.2	4.0

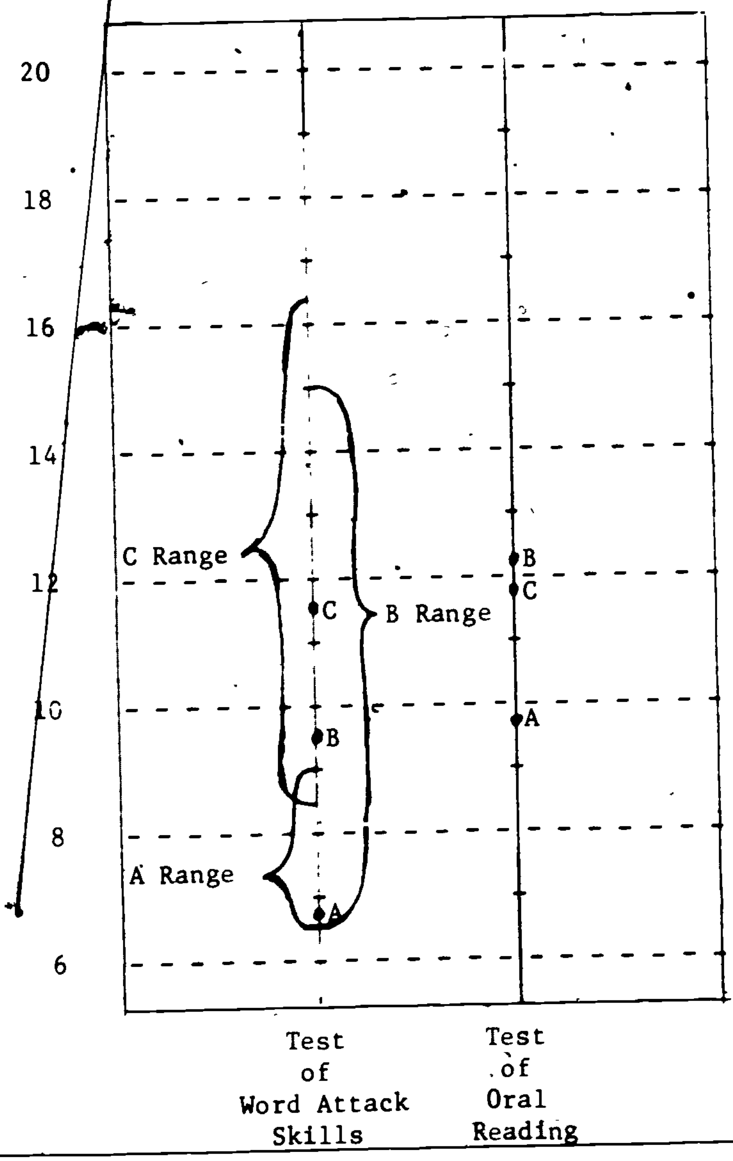
This information reinforces the discussion above regarding the difficulty of the test. Level A was obviously extremely easy. The test as a whole as well as Level B and C appeared to be relatively easy for this sample also, but there seemed to be a sufficient range in scores to indicate that the test was measuring a wide range of ability within the sample.

The Test of Oral Reading. This test was intended to consist of four passages of increasing difficulty. The first three passages were to correspond to the three parts of the Test of Word Attack Skills, Levels A through C, which also were to be of increasing difficulty. That is, Level A in the Test of Word Attack Skills was to correspond to the first passage in the Test of Oral Reading, Level B to the second passage, and Level C to the third passage; each of the three levels in the Test of Word Attack Skills were to be based upon the decoding skills defined in the hierarchy outlined by the committee and each level was to have a corresponding passage in the Test of Oral Reading. The fourth passage in this test was to be taken from the 170 reading tasks in the National Reading Performance Survey, the passage being at the fifth- or sixth-grade reading level. Figure 2.1 shows the mean level of difficulty on a common scale for each of the three parts in the Test of Word Attack Skills as well as the level of difficulty for each of the first three passages in the Test of Oral Reading. Also shown is the range of item difficulties for the Test of Word Attack Skills.

It is to be noted that for this sample the third passage (Level C) was easier than the second passage (Level B), although the percentage passing each of these was very close. The passage for Level C was passed by 60% of the sample while 59% passed the Level B passage. Eighty percent of the sample passed the Level A passage, yet this passage was more difficult than any of the items on Level A of the Test of Word Attack Skills. The level of difficulty of each of the passages was not the same as the mean level of difficulty of the part on the Test of Word Attack Skills to which it was to correspond, but the level of difficulty for each of the passages B and C was within the range of item difficulty for its corresponding part on the Test of Word Attack Skills.

The fourth passage in the Test of Oral Reading, the one from the National Reading Survey, had the highest level of difficulty. Further, the level of difficulty for this passage was higher than the mean level of difficulty of any other measure in the study. Only 32% of the sample passed this passage.

Figure 2.1. Mean Levels of Difficulty and the Range of Item Difficulties for the Levels A, B, and C on the Test of Word Attack Skills and the Difficulty of the Levels A, B, and C Passages on the Test of Oral Reading



Note: The scale used is an index of difficulty (Delta) that ranges from 6.0 to 20.0: 6.0 being very easy to 20.0 being very difficult (Thorndike, 1971, pp. 139-140).

Intercorrelations for All Measures in the Study

Table 2.8 presents the intercorrelations for the measures. Generally speaking, all but one of the measures appear to correlate with each other at a significant level. The notable exception is the Alphabet Test, a very easy test, which appears to correlate at a significant level only with the Survey of Reading Material, the total and the easy part (Level A) of the Test of Word Attack Skills. In addition, Level A of the Test of Word Attack Skills does not appear to show a significant level of correlation with the passage from the National Reading Survey, the most difficult passage in the Test of Oral Reading. It is important to note that correlation coefficients tend to be suppressed by lack of score variance, and for some measures in this study there was little score variance.

Discussion of the Results

The Decoding Instrument. In order to determine whether or not the inability to answer correctly the adult functional reading tasks in the Survey of Reading Materials is attributable to inadequate decoding skills, it is necessary to determine if the decoding instrument developed for this study was a valid and adequate measure of decoding skills. Before focusing on the decoding instrument, it is necessary to put things in perspective. The data available is based on a rather small sample, 77 adults. Further, the sample was not a random sample of the adult population consisting of those both with or without major decoding problems, although the sample selection was based on obtaining a wide range of reading ability. To this extent, conclusions drawn from the data could be rather tenuous and in need of further consideration through either replication of this study with other sample(s) or new studies which take the information here and build upon it, possibly through further development of the measures. With this in mind, it is possible to turn to a discussion of the decoding instrument.

First, the question of validity is examined. One reason for including the passages in the Test of Oral Reading was to build into the study a way of measuring concurrent validity for the paper-and-pencil test of decoding skills. Although every effort had been made to restrict the decoding skills required to read each of the first three passages in the Test of Oral Reading to only those decoding skills being measured by the corresponding part of the Test of Word Attack Skills, the data may give the impression that this was not done since the mean level of difficulty for each of the parts of the Test of Word Attack Skills does not correspond to the level of difficulty of the passage to which it was to correspond in the Test of Oral Reading. However, each test contained only a sample of the skills to be tested and it was not possible to sample exactly the same components of a skill in both the Test of Oral Reading and the Test of Word Attack Skills. Further, the data shows that the passages do not fall into a hierarchical order in terms of level of difficulty; but this may be a function of the scoring criteria. The Level B and C passages are not in the order intended in terms of level of difficulty; they are reversed. The criterion for passing the Level A or B passage was to allow no more than one error whereas for the Level C

Table 2.8

INTERCORRELATIONS FOR THE MEASURES

(N = 77)

Instrument/Test/Level	2	3	4	5	6	7	8	9	10
1. Survey of Reading Materials	.29	.64	.47	.50	.71	.46	.39	.52	.40
Decoding Instrument including:									
2. Alphabet Test		.22	.41	.19	.15	.21	.19	.17	.11
3. Test of Word Attack Skills - Total			.70	.94	.92	.46	.39	.52	.36
4. Level A				.59	.60	.49	.35	.24	.18
5. Level B					.75	.36	.29	.53	.35
6. Level C						.47	.43	.45	.33
Passages									
7. Level A							.26	.34	.27
8. Level B								.48	.29
9. Level C									.39
10. From National Reading Survey									

Note: For 75 df r's of .22 and .29 are significant at the .05 and .01 levels respectively.

passage no more than two errors were allowed. In other words, perhaps passing the Level C passage was made easier by allowing one error more to be made than was allowed for the Level A or B passage. This is certainly something to be considered in the subsequent use and/or development of this instrument. Perhaps for the reasons given above the mean levels of difficulty for the three parts of the Test of Word Attack Skills were not congruent with the levels of difficulty for the first three passages in the Test of Oral Reading, yet this does not mean that the paper-and-pencil test of decoding skills was not valid. Table 2.8 shows that the corresponding levels of the two tests are significantly correlated at the .01 level.

What becomes apparent from the data for the Test of Word Attack Skills and the Test of Oral Reading is that, given the two tests with the scoring criteria/procedures used in this study, the oral reading of the passages does not provide the range of information about a subject or a sample that the Test of Word Attack Skills does. This might not be the case, however, if the scoring were different for the oral reading. Apparently, this is an area which requires additional investigation.

Perhaps because decoding is a complex process it has been difficult to define. Even when great care has been taken in defining it, as was the case for this study, there are still questions to ask. One such question is: Does oral reading of a passage serve in determining concurrent validity for another measure that calls upon decoding skills for graphemes and morphemes not in connected discourse? Perhaps "fluency of decoding" is one factor that may affect results on one measure but not the other. And this may not be all. It might be that decoding individual graphemes and morphemes may call upon skills that are in some way different from those for decoding sentences and/or passages even when they are to be read word by word. This, of course, is not a new thought. It was recently well discussed at a conference on the relationships between speech and learning to read sponsored by the National Institute of Health (Kavanagh and Mattingly, 1972).

Before leaving the discussion of the decoding instrument, it is necessary to comment on the remaining measure, the Alphabet Test. This test was included because of the many findings that have shown knowledge of the alphabet to be a predictor of reading achievement in children (Chall, 1967, pp. 155-159). The data from the present study showed that almost the entire sample had mastered the alphabet, yet some of those who had did not do well on the other decoding measures. Further, it was noted that the Alphabet Test appeared to be significantly correlated only with the Survey of Reading Material (at the .01 level), the total Test of Word Attack Skills (at the .05 level) and Level A, the easy part of the latter test (at the .01 level); however, for the Alphabet Test there was little score variance. For another sample with greater variability in scores on this test, the correlation coefficients may be larger than they are for this sample. Speculation is possible.

Could it be that knowledge of the alphabet may not be as good a predictor for some aspects of reading achievement in adults as it is for other aspects? Perhaps it best predicts achievement for the reading skill(s) that are most like the skill(s) required for learning the alphabet, for example recognizing a symbol for its orally given referent. Clearly, additional research is needed in this area; in fact, a study including regression analysis might well provide greater insight as to the strength of a factor like knowledge of the alphabet as a predictor of the various aspects of reading achievement, granted a difficult task since the components of reading have yet to be generally agreed upon (Farr, 1969).

Ability to Decode as Related to Success on the Survey of Reading Material. As was expected, the data from this study clearly indicates that the ability to decode is a predictor of success on the Survey of Reading Material, a sample of the adult reading tasks in the National Survey of Reading. All of the decoding measures were significantly correlated with the Survey of Reading Material at the .01 level.

To illustrate this point further, Table 2.9 lists the mean scores on the Alphabet Test and the Test of Word Attack Skills for five groups of adults with increasing scores on the adult reading tasks. Each group contains approximately 20% of the sample. It is obvious from this table that those who do well on the Survey of Reading Material also tend to do well on the other tests.

Table 2.9

DISTRIBUTION OF MEAN SCORES BY QUINTILES FOR THE SURVEY OF READING MATERIALS

Survey of Reading Material			Test of Word Attack Skills			
Score Interval	Mean	Alphabet Test	Level A	Level B	Level C	Total
30-33	30.9	25.9	12.8	26.9	16.5	56.2
25-29	27.6	25.9	12.6	26.8	15.9	55.4
19-24	21.2	25.6	12.7	24.3	13.6	50.6
14-18	16.4	25.8	12.5	24.2	12.3	48.9
1-13	7.3	25.3	11.7	21.9	10.3	43.9

An attempt was made to determine a critical score on the various decoding measures which would predict success on the Survey of Reading Material (Guttman, 1941; Guilford and Michael, 1949); however, this was not possible given the data from this sample. Again, the characteristics of this particular sample and/or the scoring procedures for the various measures may have been factors in preventing the identification of the critical score. Also, the instruments appear to need further development work and refinement. Further instrument development and research is indeed indicated to determine if there is a critical decoding score (level of ability) to predict success on the adult reading tasks.

Summary

This study has been only a first step in the direction of solving the question to which it was addressed. In the discussion above, additional efforts are indicated. Clearly, the decoding instrument used needs further development work before it can be considered to be in a final form for use as a measure to diagnose decoding ability in adults. Also, further research is needed to identify a critical score or critical level of decoding ability for adult functional literacy.

REFERENCES

Chall, J. S. Learning to Read: The Great Debate. New York: McGraw-Hill, 1967.

Dale, E. and Chall, J. S. "A Formula for Predicting Readability." Educational Research Bulletin, Ohio State University, 27, 1948, pp. 11-20.

Dawson, M. A. Teaching Word Recognition Skills. Newark, Delaware: International Reading Association, 1971.

Farr, R. Reading: What Can Be Measured? Newark, Delaware: International Reading Association, 1969.

Goodman, K. S. and Fleming, J. T. (Eds.) Psycholinguistics and the Teaching of Reading. Newark, Delaware: International Reading Association, 1969.

Gray, W. S. and Leary, B. E. What Makes a Book Readable. Chicago: University of Chicago Press, 1935.

Guilford, J. P. and Michael, W. B. The Prediction of Categories from Measurements. Beverly Hills, Calif.: Sheridan Supply, 1949.

Guttman, L. The Prediction of Personal Adjustment. New York: Social Science Research, 1941.

Kavanagh, J. F. and Mattingly, I. G. (Eds.) Language by Ear and by Eye: The Relationships between Speech and Reading. Cambridge, Mass.: The MIT Press, 1972.

Kuder, G. F. and Richardson, M. W. "The Theory of the Estimation of Test Reliability." Psychometrika 2, 1937, pp. 151-160.

Mangano, J. A. (Ed.) Strategies for Adult Basic Education. Newark, Delaware: International Reading Association, 1969.

Miller, G. A. (Ed.) Linguistic Communication: Perspectives for Research. Newark, Delaware: International Reading Association, 1973.

Murphy, R. T. Adult Functional Reading Study. Final Report. Washington, D. C.: U. S. Department of Health, Education, and Welfare; National Institute of Education, 1973.

Singer, H. and Ruddell, R. B. (Eds.) Theoretical Models and Processes of Reading. Newark, Delaware: International Reading Association, 1970.

Sticht, T. G., Beck, L. J., Hauke, R. N., Kleiman, G. M. and James, J. H. Auding and Reading: A Developmental Model. Alexandria, Virginia: Human Resources Research Organization, 1974.

APPENDIX

SPECIFICATIONS FOR A RESEARCH INSTRUMENT

FOR EVALUATING WORD ATTACK SKILLS IN ADULTS

I. Naming the letters of the alphabet

II. Reading three passages which are constructed to evaluate specific skills that are believed to develop hierarchically.

The focus of each of the three paragraphs will be as follows:

*Level A - Auditory discrimination
Visual discrimination
Recognition of basic sight vocabulary
Use of context clues

**Level B - Recognition of initial consonant sounds and their graphemic options
Recognition of final (ending) consonant sounds and their graphemic options
Recognition of short vowel sounds and their graphemic options
Recognition of long vowel sounds and their graphemic options
Recognition of diphthongs and their graphemic options
Recognition of digraphs such as some vowel digraphs (ay, ea, ee) and consonant digraphs (sh, th, ch) and their phonemic options
Use of context clues
Use of structural analysis such as some derived forms (re-, dis-, mis-, un-, -est, -er, -tion, -[a]ble), pluralization, contractions

***Level C - Use of structural analysis including derived forms, pluralizations, contractions, verb tense, inflectional endings
Recognition of digraphs, including vowel digraphs, consonant digraphs, VC digraphs, r controller and their phonemic options
Use of context clues
Syllabication
Blending syllables

Note: * = Approximately upper first grade
** = Approximately upper second or early third grade
*** = Approximately middle or upper fourth grade
(A fourth paragraph from the National Reading Survey will be given following the three above (Level A-C) and this fourth paragraph will approximate a fifth or sixth-grade level).

Level A is the base level with Level B at a higher level and Level C the highest level to be tested.

III. Working through three sets of test items—one set for each passage—each of which focuses on the skills in the passage to which it pertains.

TEST ADMINISTRATION PROCEDURES

The examinee will be asked, first, to identify letters of the alphabet as they are said. Second, the examinee will be asked to read the Level A paragraph, then, the Level B paragraph and, finally, the Level C paragraph, in that order. If the examinee cannot read the Level A paragraph, he will not need to go further but will be asked if he can read any words in the other two paragraphs. If he reads the Level A paragraph, he will be asked to read the Level B paragraph. If he cannot read the Level B paragraph, he will not need to go further but will be asked if he can read any words in the Level C paragraph. If he can read the Level B paragraph, he will be asked to read the Level C paragraph.

The scoring of the paragraph readings will be as follows:

- Only 5 errors per 100 words will be permitted; this ratio of errors will be maintained for paragraphs of more or less than 100 words.

- Errors include:

  omission (except articles or words that appear more than once)
  substitution
  additions (such as insertions)
  position change
  needed aid in pronunciation
  mispronunciation (except when examinee follows rules of pronunciation but word does not)

- That which are not errors include:

  repetiton
  pausing (hesitations of short periods of time)
  not following punctuation

In addition to the paragraph reading, the examinee will be given three sets of test items covering the skills on which the paragraphs focus.

CHAPTER 3

RELATIONSHIP OF FUNCTIONAL READING TO PERFORMANCE ON CLOZE TESTS

In Chapter 2, the relationships between functional reading and decoding skills were studied in order to determine whether inability to read functional materials is due to lack of decoding skills. In this chapter, relationships between functional reading and performance on cloze tests are examined. The cloze test, as explained below, is a method for assessing reading skill. Measures obtained by cloze tests have been related to measures obtained by more traditional tests in an attempt to validate the method. In this phase of the study, scores on functional reading tests and scores on cloze tests are related in an attempt to demonstrate concurrent validity between these two methods of measuring reading skill.

Selected Reading Tasks

The reading tasks were a subset of the 170 functional reading tasks developed by Educational Testing Service as part of the Adult Functional Reading Study. Each reading task consists of two components: 1) a stimulus that is a sample of realistic reading material the comprehension of which is judged to be of high benefit to the reader; and 2) a required performance that serves to indicate whether or not the stimulus was comprehended. The reading tasks have been administered to a national probability sample of 8,000 adults and data on the difficulty of each task are available (Murphy, 1973).

Selected Cloze Tests

The cloze tests were based on a subset of 36 passages that had been scaled for difficulty using an extensive set of tests (Coleman and Miller, 1968). Available for each passage are data related to difficulty based on: 1) percentage of correct responses on cloze tests in which every fifth word was deleted; 2) percentage of correct responses on cloze tests in which one word per passage was deleted; 3) percentage of correct responses when subjects guessed every word after seeing the preceding word in the passage; 4) percentage of correct responses when subjects guessed every word after seeing the preceding word, after having read the passage once; 5) the difference between #3 and #4 which is interpreted as information gain; 6) mean number of content words an adult recalls after studying the passage for 60 seconds; 7) mean subjective rank order of difficulty as ranked by 14 judges (Aquino, 1969). It should be pointed out that these data, while extensive, are based on college student subjects, and the degree to which they are generalizable to other populations remains to be established.

Advantages of Cloze Tests

Cloze tests are of interest for both practical and theoretical reasons. The primary practical advantage is that cloze tests are exceedingly simple to construct. Every nth word is deleted from an appropriate passage and the subject is asked to "fill in the blanks." That exceedingly simple test construction process leads to scores that are "highly valid measures of the readability of printed materials and the comprehension of readers" (Rankin and Culbane, 1969).

From a theoretical point of view, cloze tests have a significant advantage over traditional tests of reading comprehension in that the construction of cloze tests can be based on an algorithmic, replicable procedure free of bias. Consider the fact that traditional reading tests generate scores based on the questions that have been written to measure comprehension of a particular passage. It is well known that questions of various levels of difficulty may be written for the same passage. For multiple-choice questions, the same stem will provide items of various levels of difficulty depending on the distracters that are selected to accompany that stem. For a given population, the percentage of correct responses to a set of questions on a reading passage depends not only on the inherent difficulty of the passage itself, but also on the particular set of question stems that happen to have been written and on the particular set of distracters chosen for each question stem. The choice of question stems and distracters is, to a great extent, a function of the personal biases of the item writer.

Using traditional tests of reading comprehension, it is not possible to make rigorous judgments of the difficulty of written text. Measures of the difficulty of a set of questions based on a given text are obtained and inferences are made about the difficulty of the text. It must be kept in mind, however, that a different set of questions based on the same text would lead to different judgments of the difficulty of that text. On the other hand, it must be noted that traditional reading tests allow the measurement of relatively distinct skills such as understanding of the main idea, making inferences, etc., which are not measured by cloze tests. While the question of the exact nature of the skill tapped by the cloze procedure remains the subject of debate, and the way in which this skill is related to the reading process requires further elucidation, it is clear that success in this task requires a sensitivity to the redundancy and syntactic structure of English. Thus cloze tests go beyond the vocabulary and general knowledge demands of traditional comprehension items, and are compatible with linguistic theories of reading.

The use of cloze tests obviates the biasing effect of the item writer. It becomes possible to strictly delimit the universe of all possible cloze items for any given string of text and to generate strictly random samples from that universe allowing unbiased estimates of performance on the universe of items from performance on the sample of items (Bormuth, 1974).

Score Interpretation of Cloze Tests

Because the use of cloze tests has been rare in comparison with the use of traditional measures of reading ability, potential users have been unable to generate an intuitive understanding of various levels of cloze test performance. Is 35 percent correct an acceptable score? Or is 85 percent correct required to indicate comprehension of reading material has been attained?

The wide experience with traditional measures has shown certain "rules of thumb" to be useful: 75 percent correct on a set of traditional questions has been generally accepted as indicating a level of difficulty that is suitable for instructional use, and 90 percent correct has been accepted as indicating a level of difficulty that is suitable for independent study (Harris, 1962). This, in spite of the fact that a given absolute percent is a function of the items to at least as great an extent as of the difficulty of the passage.

Efforts have been made to show the relationship between cloze test scores and multiple-choice test scores (Bormuth, 1967; Rankin and Culhane, 1969). In general, a cloze score of 40 ± 4 percent can be taken to be equivalent to a raw score of 75 percent on a typical multiple-choice test based on the same reading materials.

Comparable Reading Task and Cloze Test Difficulties

One way of relating the reading tasks and the cloze passages was to compare the difficulty levels of the selected reading tasks with the difficulty levels of the selected cloze tests. Each reading task produced a binary score: right or wrong. Each cloze test, however, could produce a range of scores from 0 to 30, scoring one point for each deletion correctly replaced. The difficulty of each reading task could easily be stated as the percent of the population that selected the correct response. In order to obtain comparable data for each cloze passage, a cutting point had to be selected on the score scale at or below which the passage would be scored as "wrong," and above which the passage would be scored as "right."

Bormuth (1971) has given evidence that readers who score less than approximately 35 percent on cloze tests can gain little or no information from materials at that level of difficulty. Since the reading tasks were designed primarily to discriminate between literate and illiterate populations, the score level congruent with the ability to gain information from a passage was selected as the cutting point. For the current study, the difficulty of a cloze passage can, therefore, be stated in terms of the percent of the population that score 11 or more deletions correctly replaced. That percentage can be compared directly with the percentage of the population responding correctly to a reading task.

Clearly, the mapping of reading task difficulty into cloze test difficulty is not unique. It will vary if the cutting score on the cloze test varies. For that reason, data are reported that will allow the mapping to be done for all possible cutting scores. Such a comparison is meaningful only to the extent that the two sets of tasks covary. That is, to the degree that those who pass the reading tasks include the people who experience success on the cloze tasks. The extent to which this is the case will be discussed in a later section.

Test Description

Six tests were constructed, each containing 18 reading tasks and three cloze passages. The three cloze passages and three reading tasks, common across grades, were placed in the same position in each test booklet. The 15 remaining reading tasks were ranked according to difficulty level on the national survey. However, rather than simply ordering the tasks in the test booklets according to difficulty level, four of the easier tasks were placed at the end of each test booklet. This was done to insure that students would have the opportunity to attempt the more difficult reading tasks. Our intention was to allow sufficient time so that non-responses could be interpreted as inability to respond correctly. If some students did run out of time, then the last few items could be deleted from the analysis without losing the range of item difficulties present in the study.

A complete description of the six tests, two for each of grades 7, 10, and 12 is given in Table 3.1. Cloze passages are identified by the letters A through H. The reading tasks are identified by the listings used in the national survey. Reading tasks administered across grades are noted by asterisks.

Table 3.1

TEST DESCRIPTION

Page	Grade 7		Grade 10		Grade 12	
	I	II	I	II	I	II
1	4-2	4-3	5-6	7-14	6-1	9-11
2	2-2	1-2	6-8	10-4	1-17	5-10
3	A	B	B	C	C	D
4	2-3	4-1	3-10	10-9	4-10	1-16
5	C	D	D	E	E	F
* 6	1-1	8-1	1-1	8-1	1-1	8-1
7	7-2	1-3	6-2	5-1	6-14	8-8
8	3-1	2-1	5-16	6-6	2-14	2-17
9	E	F	F	G	G	H
10	6-3	2-10	2-8	6-12	1-10	5-11
*11	10-11	7-3	10-11	7-3	10-11	7-3
12	7-5	8-4	10-7	3-12	8-13	10-15
13	8-6	9-7	7-6	9-6	4-15	4-13
14	4-6	3-4	1-11	7-4	3-13	8-9
*15	6-9	6-13	6-9	6-13	6-9	6-13
16	10-10	9-5	8-5	7-12	3-15	10-17
17	1-14	1-6	9-9	3-5	5-15	9-13
18	2-6	9-2	3-8	8-7	8-15	6-16
19	9-1	5-2	1-7	10-2	8-12	4-16
20	10-8	5-3	7-9	6-4	6-11	9-10
21	2-11	9-14	5-13	4-9	2-16	9-16

Test Administration

The tests were administered by four ETS staff members in student classrooms. Each class was tested on two consecutive days. Students who were absent on one of the days completed only one test. Average administration time was approximately 45 minutes. Some teachers remained in the classroom, assisted in distributing the materials, and showed considerable interest in the test administration and materials.

After the tests were distributed, brief instructions were given to the students on marking procedures for both the reading tasks and the cloze passages. Students indicated their responses directly in the test booklets. Although the test was not strictly timed, the students were asked to move on to pages 4, 6, and 10 after intervals of approximately ten minutes if they had not already done so. The students were encouraged to go back over the test if they finished early.

Scoring

Each student was assigned a unique identification number which was used to match the two test booklets for the two consecutive days. Then the reading tasks were scored as right, wrong, or omit using the coding specifications for the national survey. The cloze passages were scored allowing exact replacement only. Every item was scored as right, wrong, or omit. Interpretable misspellings were scored as correct. The item scores were then transfered to machine scorable answer sheets for the analysis.

Results

Based on the observations of the ETS test administrators, most students found the materials to be interesting and motivating, although some 12th grade basic classes showed a disposition to treat the task with no more seriousness than, from their perspective, it deserved. Some students asked questions about the materials at the end of the testing session; a few were observed discussing the materials and arguing about their responses. The total sample of students tested was 787: 244 in grade 7, 257 in grade 10, and 286 in grade 12. In all, 32 classes in 7 schools participated in this study.

The means and standard deviations for the six functional reading tests and the eight cloze tests are given in Table 3.2. The reading tasks administered to the 12th grade students were more difficult than those administered to the 10th grade students. The least difficult tasks were administered to the 7th grade students. The complete item statistics for the reading tasks are given in Table 3.3. As explained earlier, it was projected that students would have sufficient time to complete all the tasks and that non-responses would be interpreted as inability to respond correctly. Moderately large percentages of students omitted the last few items. For this reason, the percentage of non-responses is also given for each item. For comparison purposes, the

Table 3.2

DESCRIPTIVE STATISTICS

	Grade 7 Mean (SD)	Grade 10 Mean (SD)	Grade 12 Mean (SD)
Reading Tasks			
Test I	14.8 (4.2)	13.9 (4.8)	10.0 (5.3)
Test II	14.7 (4.3)	13.9 (5.2)	10.9 (5.8)
Cloze Tests			
A	13.7 (6.3)		
B	17.9 (5.8)	17.0 (6.0)	
C	16.3 (6.5)	17.5 (6.9)	16.3 (7.5)
D	12.5 (5.7)	13.7 (5.8)	12.6 (7.3)
E	5.3 (4.0)	6.8 (4.6)	6.8 (5.1)
F	7.9 (4.6)	9.2 (5.2)	9.2 (6.0)
G		8.7 (5.4)	8.7 (5.7)
H			12.4 (13.2)

difficulty levels from the national survey are also included in Table 3.3. Common items have been marked with asterisks to allow comparison across grades. In general, the 10th grade students performed better than the 7th grade students, and the 12th grade students performed about as well as the 10th grade students.

The percentages of students who "passed" the cloze passages, i.e. filled in 11 or more blanks correctly, are given in Table 3.4. On the more difficult cloze passages, the students performed as would be expected. On passage B (grade level 1.5), the 10th grade students scored slightly lower than the 7th grade students. On passage C (grade level 5.7), the 12th grade students performed less well than the 10th grade students. This is not an uncommon finding when a single passage is used for a cloze test across grades. When the vocabulary included in the passage is more akin to that of the younger students, older students tend to use more complex synonyms. Since scores are based on exact word replacement, older students tend to score less well. The result is not important in this study as the relationships between the cloze tests and the reading tasks are interpreted only within grades. That is to say, a cloze test and a reading task may have similar difficulty levels for grade 12. That relationship may or may not hold true for the other grades. For example, passage C is about as difficult for the 12th grade students as is the common reading task 10-11. That is not the case for the 7th grade students. Table 3.4 also presents coefficient alpha reliabilities for each passage. As expected, maximum reliability occurs for each grade if a passage is of intermediate difficulty for that grade.

Table 3.3

READING TASKS - ITEM STATISTICS

GRADE 7

Test I					Test II				
Task	% Pass Nat'l Survey	% Pass Omit = Wrong	% Omits	r-biserial	Task	% Pass Nat'l Survey	% Pass Omit = Wrong	% Omits	r-biserial
4-2	100	99	1	0.79	4-3	99	90	6	0.67
2-2	99	98	1	0.68	1-2	99	97	0	*
2-3	98	94	2	0.86	4-1	98	100	0	0.99
*1-1	97	89	0	0.64	*8-1	90	76	2	0.80
7-2	98	90	4	0.96	1-3	97	94	1	*
3-1	96	91	1	0.91	2-1	95	93	1	*
6-3	95	93	1	0.70	2-10	94	88	3	0.80
*10-11	85	71	3	0.80	*7-3	76	79	7	0.75
7-5	94	82	3	0.61	8-4	93	90	3	*
8-6	93	88	1	0.52	9-7	93	87	6	*
4-6	92	87	4	*	3-4	91	80	4	0.77
*6-9	73	52	6	0.66	*6-13	60	32	2	0.46
10-10	91	86	5	*	9-5	90	89	3	0.92
1-14	90	83	6	0.93	1-6	89	77	4	0.81
2-6	98	91	6	*	9-2	96	89	5	*
9-1	96	89	7	*	5-2	94	90	4	*
10-8	93	78	10	0.96	5-3	93	87	3	*
2-11	92	80	11	0.92	9-14	89	80	4	0.92

*The r-biserial was greater than 1 due to lack of normality in the dichotomized variable and the inexact nature of the estimation procedure.

Table 3.3 (con't)

READING TASKS - ITEM STATISTICS

GRADE 10

Test I					Test II				
Task	% Pass Nat'l Survey	% Pass Omit = Wrong	% Omits	r-biserial	Task	% Pass Nat'l Survey	% Pass Omit = Wrong	% Omits	r-biserial
5-6	87	83	5 5	0.53	7-14	86	81	8	0.70
6-8	86	72	11	0.66	10-4	86	89	2	0.47
3-10	86	84	3	0.76	10-9	85	95	1	0.88
*1-1	97	98	1	0.31	*8-1	90	93	1	0.85
6-2	85	93	2	0.76	5-1	85	91	1	0.58
5-16	85	84	4	0.77	6-6	84	86	1	0.61
2-8	83	69	5	0.81	6-12	82	91	3	0.89
*10-11	85	89	3	0.77	*7-3	76	86	5	0.66
10-7	82	83	2	0.64	3-12	81	84	8	*
7-6	81	81	2	0.69	9-6	81	91	2	0.68
1-11	80	74	7	0.95	7-4	79	89	3	0.82
*6-9	73	69	3	0.85	*6-13	60	58	3	0.68
8-5	79	80	11	0.93	7-12	78	78	8	0.85
9-9	78	88	8	0.83	3-5	76	82	5	0.88
3-8	88	88	10	0.90	8-7	87	88	4	0.83
1-7	86	75	12	0.89	10-2	84	89	5	0.83
7-9	83	79	13	0.98	6-4	83	79	5	0.67
5-13	82	75	20	0.88	4-9	81	78	6	0.89

*The r-biserial was greater than 1 due to lack of normality in the dichotomized variable and the inexact nature of the estimation procedure.

Table 3.3 (con't)

READING TASKS - ITEM STATISTICS

GRADE 12

	Test I					Test II			
Task	% Pass Nat'l Survey	% Pass Omit = Wrong	% Omits	r-biserial	Task	% Pass Nat'l Survey	% Pass Omit= Wrong	% Omits	r-biserial
6-1	76	64	18	0.47	9-11	75	90	2	0.65
1-17	74	67	17	0.58	5-10	74	84	2	0.54
4-10	73	87	5	0.86	1-16	72	79	3	0.59
*1-1	97	97	1	0.63	*8-1	90	95	1	0.41
6-14	71	82	2	0.80	8-8	71	91	1	0.78
2-14	70	82	4	0.38	2-17	69	88	3	0.54
1-10	67	77	4	0.42	5-11	64	70	4	0.58
*10-11	85	93	5	0.66	*7-3	76	92	4	0.63
8-13	63	54	7	0.45	10-15	52	79	4	0.77
4-15	50	60	3	0.51	4-13	48	65	3	0.61
3-13	47	58	4	0.67	8-9	47	62	3	0.62
*6-9	73	70	22	0.87	*6-13	60	65	8	0.61
3-15	44	31	32	0.73	10-17	44	65	12	0.68
5-15	35	27	44	0.81	9-13	34	38	13	0.56
8-15	75	58	38	0.87	6-16	71	70	21	0.86
8-12	71	50	46	0.92	4-16	68	60	30	0.75
6-11	65	52	46	0.83	9-10	63	57	28	0.77
2-16	58	42	51	0.85	9-16	54	54	33	0.71

* The r-biserial was greater than 1 due to lack of normality in the dichotomized variable and the inexact nature of the estimation procedure.

Table 3.4

CLOZE TESTS
ITEM STATISTICS

Cloze Passage	Grade Level	Grade 7 % Pass	alpha	Grade 10 % Pass	alpha	Grade 12 % Pass	alpha
A	7.1	77	(.87)				
B	1.5	93	(.81)	92	(.70)		
C	5.7	88	(.87)	96	(.79)	88	(.77)
D	6.0	78	(.86)	84	(.83)	85	(.84)
E	6.8	9	(.81)	26	(.83)	31	(.84)
F	5.9	29	(.83)	51	(.89)	63	(.83)
G	9.6			49	(.85)	53	(.85)
H	7.9					22	(.83)

As mentioned at the beginning of the chapter, the criterion score used on the cloze tests was 11. Students who filled in fewer than 11 blanks correctly failed the cloze test. Other cutting scores are possible. Complete distributions of the students on the cloze tests are available for every combination of reading task and cloze passage. These data are for those students who responded to both. It allows cloze passages and reading tasks to be compared across days. Of course, the numbers of students included for the various combinations are different. For this reason, the difficulty levels for the cloze passages and the reading tasks differ from the overall values given in Tables 3.2 and 3.3. However, this information is noteworthy. Many students who fail the cloze passage manage to pass many of the reading tasks. This is another way of looking at the difficulty levels. In view of the fact that the stated grade levels for the cloze passages ranged from 1.5 to 9.6, the data suggest that many students who are unable to "gain information" from these passages are able to respond correctly to ordinary real life reading tasks.

Relationships Between the Cloze Tests and the Reading Tasks

Why should any relationship be expected between these cloze tests and the reading tasks? Both are tests of reading skills. Admittedly, they may relate to different facets of reading. The reading tasks are everyday kinds of tasks. The cloze passages are paragraphs from a series of reading texts. In general, the scores of the students on these two tests are moderately correlated. Students who perform well on one test perform well on the other , in general. Thus the two tests tap a common factor of ability as well as possessing specific and error variance.

Table 3.5 gives intercorrelations of the total reading and cloze scores for each day. Correlations above the main diagonal are based on all students present for both days of testing. Students who omitted entire cloze passages were given zero scores for those tests. The generally lower correlations below the main diagonal are missing data correlations, with students not attempting certain cloze passages counted as missing for those passages. Entries on the main diagonal are reliabilities, coefficient alpha in the case of the reading tasks, and split-third reliabilities based on each day's median correlations for the cloze passages.

Examination of this table reveals that the reliabilities range from .78 to .90, with the lowest reliabilities occuring among the 12th grade students. The highest correlations in each grade are those of the two days' cloze totals. The two sets of reading tasks were next most strongly related, and tended to covary with the cloze scores at almost as high a level, with a value of about .70 for grades 7 and 10, and .55 for grade 12. Interestingly, no consistent evidence emerged of a higher relationship among reading tasks and cloze tests given on the same day than in those given on different days. For each grade, reliabilities for the total reading and cloze scores are also given in Table 3.5, along with the correlations of these totals. It is clear that the two tasks have a large proportion of their reliable variance in common. In grade 10, for example, the reliabilities of the two instruments would place a theoretical upper limit of .87 on their intercorrelation if they were measures of exactly the same trait. The observed correlation, .77, tells us that nearly 60% out of a possible 78% of the variance in the cloze tasks is predictable by reading task performance. While not unidimensional, the tasks seem to be tapping quite similar abilities.

The above results indicate that it makes sense to compare the difficulty levels and at least propose tentatively that students who can complete certain passages can probably pass certain common functional reading tasks. In grade 7, for example, passage A, using the 35% criterion score, is about as difficult as items 10-11 and 7-3. If the items are ranked according to difficulty level, then specific passages can be matched at appropriate levels. This has been done in Table 3.6. Because the reading tasks were generally easier than the cloze passages, in some instances a cloze passage ranks below all the reading tasks. If students could read such a cloze passage they could probably read most of the reading tasks above it. Of course, it does not follow that students who cannot read the passages cannot read the less difficult reading tasks.

Table 3.5

RELIABILITIES AND INTERCORRELATIONS

Grade 7

	R I	R II	Cl I	Cl II
R I	(.84)	.76	.73	.70
R II	.76	(.86)	.71	.70
Cl I	.65	.63	(.90)	.87
Cl II	.65	.68	.81	(.84)

Cl I = A + C + E $\quad r_{CR} = .78 \quad r_R = .86$

Cl II= B + D + F $\quad r_C = .93$

Note: Above the diagonal, N = 227. Below the diagonal, N ranges from 196 to 227.

Grade 10

	R I	R II	Cl I	Cl II
R I	(.85)	.71	.71	.72
R II	.71	(.81)	.63	.67
Cl I	.75	.66	(.84)	.84
Cl II	.71	.64	.79	(.84)

Cl I = B + D + F $\quad r_{CR} = .77 \quad r_R = .83$

Cl II= C + E + G $\quad r_C = .91$

Note: Above the diagonal, N = 223. Below the diagonal, N ranges from 200 to 223.

Grade 12

	R I	R II	Cl I	Cl II
R I	(.83)	.67	.63	.51
R II	.67	(.79)	.62	.52
Cl I	.61	.58	(.78)	.68
Cl II	.47	.46	.61	(.80)

Cl I = C + E + G $\quad r_{CR} = .68 \quad r_R = .80$

Cl II= D + F + H $\quad r_C = .81$

Note: Above the diagonal, N = 202. Below the diagonal, N ranges from 162 to 202.

---

r_{CR} = correlation of total reading and cloze scores

r_R = reliability of total reading score

r_C = reliability of total cloze score

R I: First day's total reading tasks score
R II: Second day's total reading tasks score
Cl I: First day's total cloze passage scores
Cl II: Second day's total cloze passage scores

Table 3.6

RANKING OF DIFFICULTY LEVELS
FOR READING TASKS AND CLOZE PASSAGES

GRADE 7				GRADE 10				GRADE 12			
Task	Test I	Task	Test II	Task	Test I	Task	Test II	Task	Test I	Task	Test II
4-2	99	4-1	100	*1-1	98	C	96	*1-1	97	*8-1	95
2-2	98	1-2	97	6-2	93	10-9	95	*10-11	93	*7-3	92
2-3	94	1-3	94	B	92	*8-1	93	C	88	8-8	91
6-3	93	2-1	93	*10-11	89	5-1	91	4-10	87	9-11	90
3-1	91	B	93	9-9	88	6-12	91	6-14	82	2-17	88
2-6	91	4-3	90	3-8	88	9-6	91	2-14	82	D	85
7-2	90	8-4	90	3-10	84	10-4	89	1-10	77	5-10	84
*1-1	89	5-2	90	5-16	84	7-4	89	*6-9	70	1-16	79
9-1	89	9-5	89	D	84	10-2	89	1-17	67	10-15	79
8-6	88	9-2	89	5-6	83	8-7	88	6-1	64	5-11	70
C	88	2-10	88	10-7	83	6-6	86	4-15	60	6-16	70
4-6	87	9-7	87	7-6	81	*7-3	86	3-13	58	4-13	65
10-10	86	5-3	87	8-5	80	3-12	84	8-15	58	*6-13	65
1-14	83	3-4	80	7-9	79	3-5	82	8-13	54	10-17	65
7-5	82	9-14	80	1-7	75	7-14	81	G	53	F	63
2-11	80	*7-3	79	5-13	75	6-4	79	6-11	62	8-9	62
10-8	78	D	78	1-11	74	7-12	78	8-12	50	4-16	60
A	77	1-6	77	6-8	72	4-9	78	2-16	42	9-10	57
*10-11	71	*8-1	76	2-8	69	*6-13	58	3-15	31	9-16	54
*6-9	52	*6-13	32	*6-9	69	G	49	E	31	9-13	38
E	9	F	29	F	51	E	26	5-15	27	H	22

At each grade level, students who failed to meet the 35% criterion on cloze passages of moderate difficulty managed to pass a large number of adult functional reading tasks. For example, in grade 10, item 1 reproduces a soup label containing the brand name, contents, directions, a recipe, and the ingredients which are labeled "prepared from." The item asks the students to circle the information that tells "what the soup is made from." On one level, this may seem to be a vocabulary item to test understanding of the word "prepared." Of the 231 students with complete data, 201, or 86.6% answered this item correctly. Cloze passage D, a set of directions for making a kite with a putative grade level of 6.0, was passed by 84.8% of the students. Of the 196 students who passed the cloze passage, 178 or 91% also passed the reading task. Of the 35 who failed the cloze passage, 24 or 68% also passed the reading task. This implies that the students who fail the "literacy criterion" on this particular passage are still able to respond correctly to many of the everyday functional reading tasks.

Summary

In this phase of the study, we have related the functional reading tasks developed in the Adult Functional Reading Study to a few cloze tests. In general, the functional reading tests were considerably less difficult than the cloze tests. Only eight cloze tests were used, which were available from previous research. The relationships found seem reasonable and suggest further studies with materials more appropriate to given grade levels. Especially interesting would be a study of cloze tests developed on more functional kinds of reading passages.

REFERENCES

Aquino, M. "Validity of Miller-Coleman Readability Scale," Reading Research Quarterly, IV, 3, 1969.

Bormuth, J. "Comparable Cloze and Multiple-Choice Comprehension Test Scores," Journal of Reading, 10, 1967, 291-299.

Bormuth, J. Development of Readability Analyses, University of Chicago, 1969.

Bormuth, J. "Reading Literacy: Its Definition and Assessment," Reading Research Quarterly, IX, 3, 1973-1974.

Coleman, E., and Miller, G. "A Measure of Information Gained During Prose Learning," Reading Research Quarterly, III, 3, 1968, 369-385.

Harris, A. Effective Teaching of Reading, New York: McKay, 1962.

Murphy, R. T. Adult Functional Reading Study. Final Report. Washington, D. C.: U. S. Department of Health, Education, and Welfare; National Institute of Education, 1973.

Rankin, E., and Culhane, J. "Comparable Cloze and Multiple-Choice Comprehension Test Scores," Journal of Reading, December, 1969, 192-198.

CHAPTER 4

BASIC READING COMPETENCY IN THE SCHOOLS

Are the reading tasks developed in the Adult Functional Reading Study suitable for assessing the reading skills of school students? In an attempt to answer this question, ETS entered into a joint research venture with the Bureau of Elementary and Secondary Educational Testing of the New York State Department of Education. ETS agreed to work with staff members of New York State in developing a preliminary test for experimental use with a sample of ninth grade students throughout the state.

Instrument Development

Staff members of the Bureau of Reading of the New York State Department of Education reviewed the reading tasks available in the Adult Functional Reading Study for appropriateness for use at the secondary level and for interest to students. One hundred and twenty reading tasks were selected for trial administration. These tasks were in multiple-choice format as used in the Group Administration Phase of the ETS project or were adapted to multiple-choice format from the national survey formats. Consultants with experience in the field of reading at the secondary level were invited to review the selected tasks. Revisions and deletions were made on the basis of these reviews. The tasks were then grouped into eight forms with 14 or 15 reading tasks in each form. In developing the eight forms, a moderate balance was built in based on the following criteria: topic, skill, difficulty level, benefit, and form.

First Field Testing

A sample of schools was randomly selected from a population of New York State public schools with the following restrictions:

1. that the sample be representative according to community type, i.e., New York City, large city, small city, village-suburban, and rural.

2. that the sample be representative, by community type, of the percentage of students receiving scores falling below the Statewide Reference Points on the 1973 Reading PEP test.

A summary of the sample of 2,033 ninth grade students is given below:

School	Community Type	No. of Students
#1	New York City	245
#2	New York City	414
#3	Large City	176
#4	Small City	174
#5	Village - Suburban	230
#6	Village - Suburban	206
#7	Village - Suburban	307
#8	Rural	159
#9	Rural	122

Students were allowed approximately 45 minutes to complete a form (14-15 items). This test was carried out in May and June, 1974.

Results

Results show that the difficulty levels for the items ranged from .35 to .97. The average difficulty for the eight forms was .80. A distribution of item difficulties in the eight forms is given in Table 4.1.

Table 4.1

Difficulty Level	No. of Items	Percent of Total Items
.90+	48	41
.80-.89	26	22
.70-.79	14	12
.60-.69	12	10
.50-.59	8	7
.40-.49	5	4
.30-.39	3	3

In general, the items discriminated fairly consistently between upper and lower scoring halves of the students. Items with poor discrimination indices were examined for possible revision. The tasks were then grouped into two sets of 40 items each with some items being deleted and some revised. The two sets of 40 tasks each were then used as two forms of a new experimental "test."

Second Field Testing

In November, 1974, the two forms of the experimental "test" were administered to all ninth grade students in eleven schools; each student received one of the two forms. The sample of schools used in this second field testing was randomly selected from a population of New York State public schools, with the following restrictions:

1. that the sample be representative according to community-type, i.e., New York City, large city, small city, village - suburban, and rural. The schools were scattered throughout the state.

2. that the sample include an intentionally greater percentage of pupils falling below the Statewide Reference Point on the 1973 Reading PEP test. Specifically, the target sample included approximately 40% of the pupils falling below that particular standard as compared to the 30% that actually did fall below the Statewide Reference Point.

A summary of the sample of 2,520 students is given below:

School	Community Type	No. of Students	% of Pupils Below SRP on 73 Reading PEP Test
A	New York City	485	37
B	Rural	160	37
C	Village - Suburban	375	38
D	Village - Suburban	375	41
E	Rural	155	41
F	Small City	195	41
G	Small City	275	37
H	Large City	550	38

Students were allowed one full class period (40-45 minutes) to complete a test. The tests were administered by classroom teachers. Responses were indicated on separate answer sheets. Several members of the ETS project staff and the New York State Bureau of Reading observed the test administration in two schools and interviewed a small number of teachers and students to get their reactions. In general, the reactions to these materials were favorable. Teachers thought the materials were reasonable for assessing functional reading. Students thought the materials were interesting and were motivated to respond to them.

Results

Approximately 1,000 students responded to each form of the experimental test. The total possible score on each form was 40. The mean score on form A was 30.38 (S.D. = 6.46); the mean score on form B was 30.10 (S.D. = 6.48). The difficulty levels of the items in form A ranged from .34 to .95 with a mean item difficulty level of .76. On form B, difficulty levels ranged from .33 to .96 with a mean item difficulty level of .78. A distribution of item difficulties in each form is given in Table 4.2. Using the Kuder-Richardson Formula 20, the reliability of form A was .85 (S.E. = 2.5). KR20 for form B was .84 (S.E. = 2.5). In general, the number of students omitting items increased rapidly toward the end of the test. The test was not intended to be a speeded test. Therefore, this information indicates that the test should be shortened or the time allowed for administration lengthened.

Table 4.2

Difficulty Level	Form A # of Items	Form A % of Total Items	Form B # of Items	Form B % of Total Items
.90+	10	25	13	33
.80-.89	13	33	9	23
.70-.79	7	17	5	13
.60-.69	0	0	6	15
.50-.59	5	13	3	7
.40-.49	2	5	3	7
.30-.39	3	7	1	3

For comparison purposes, the scores of the students on the 1973 Reading PEP test were also collected. The mean score for students who responded to form A was 32.62 (S.D. = 8.76) on the PEP test. The mean score for students who responded to form B was 32.58 (S.D. = 8.61) on the PEP test. The total possible score on the PEP test was 50. The correlations of PEP scores with scores on the experimental basic reading competency test was .676 and .657 on forms A and B respectively. The percentages of students falling below the Statewide Reference Point on the PEP test were 37% and 30% respectively for the two groups. If 65% of the items on the basic reading competency test is used as an arbitrary cutoff score for comparison purposes, then approximately 20% of the students are below this cutoff score. Whether such an arbitrary score can be assigned real meaning is a question that must be seriously considered by persons who use this information. The reading tasks in the experimental test have a certain face validity. But whether or not students who can respond correctly to 25 out of 40 such items are in some way sufficiently prepared for surviving in a real life reading world is an unresolved question at present. Nevertheless, this kind of information has not been collected heretofore. It may be helpful to decision makers in setting goals, assigning priorities, and assessing progress if reasonable agreement can be reached that the test measures something valuable for the educational system in general and for students in particular. This, as yet, has not been done.

The correlations (.676 and .657) reported in the preceding paragraph indicate that, in general, students who perform better than other students on the PEP tests also perform better on the experimental test. However, an alternative way of looking at the relationship between performance on the PEP test and performance on the basic reading competency test might be useful. New York State commonly uses stanines for identifying groups of students. In general, students are ordered from low to high on some test score. Then, the scores are grouped with the following percentages in each of nine groups from 1 to 9: 4%, 7%, 12%, 17%, 20%, 17%, 12%, 7%, and 4%. Thus, stanines 1-3 contain 23% of the students, 4-6 contain 54% of the students, and 7-9 contain 23% of the students. Those students in the lowest three stanines are considered to be below some cutoff point. The score corresponding to the dividing point between the third and fourth stanines is the cutoff score. This score is used as a Statewide Reference Point. If the division into stanines is performed each year, then a new Statewide Reference Point would be determined each year, and there would always be 23% of the students below the reference point. More commonly, a reference point is determined for a given year and then retained for several years. Then progress can be verified if the percentage of students scoring below the reference point decreases. In New York State, a Statewide Reference Point was established in 1966 according to this procedure. Approximately 30% of the students were below this cutoff point in 1973. In the sample of students who responded to the basic reading competency test, approximately 38% were below the cutoff point. Fixing the reference point is arbitrary. But it at least sets a bench mark against which progress can be assessed. For comparison purposes, the percentages of students who fell above and below the arbitrary cutoff score of 25 on the basic competency test within the three groups of stanines 1-3, 4-6, and 7-9 is given below:

	Stanine		
	1-3	4-6	7-9
Scored 25 or better	21%	44%	15%
Scored below 25	17%	3%	0%

Note that in this biased sample of students 38% had scored below the Statewide Reference Point on the PEP test. Approximately 55% of those students "passed" the basic reading competency test. For students above the Statewide Reference Point, about 5% of those students "failed" the basic reading competency test. The main point of this illustration is that a somewhat different set of students may be identified as in some way lacking in reading skills. Due to the nature of the materials in the basic reading competency test, it may then be more possible to explain to students the kinds of reading tasks that they are unable to handle well than is the case with some standardized tests. In the long run, students might be better motivated to improve their reading skills if such information is available to them.

CHAPTER 5

THE ECONOMIC BENEFITS OF SCHOOLING AND READING COMPETENCE*

Our purpose in this task is to estimate a statistical model of the effects of schooling and literacy on earnings and employment that will allow us to calculate the economic benefits of schooling and compensatory education. Our approach is in the spirit of the earnings function literature, which has been recently and thoroughly reviewed by Psacharopoulos (1973), but our findings extend the existing literature in three important ways. First, our data are from a 1972 probability sample of the U.S. population over the age of 16; previous earnings functions have been for restricted subgroups of the population, usually white males, and our data, therefore, allow more detailed analysis of the effects of race and sex on earnings than has hitherto been possible within a single data set. Second, our data allow construction of a block-recursive model that examines first the interaction of education and reading skill and, second, the determinants of wage rate and labor supply. The economic benefits of schooling and literacy skills can thus be decomposed into not only their effects on each other, but also on wage rate and labor supply. Third, and most important, our data set includes a measure of each respondent's basic reading capability. The measure was derived from a set of reading tasks constructed to assess the respondent's capability to read the kind of material that appears frequently in day-to day life in the contemporary United States; it thus differs in important ways from the ability measures that appear in a number of earnings functions. Perhaps its most important difference is that the imparting of reading competence at this level is perhaps the foremost single objective stated by public school systems; and over the last decade there has been a massive national effort (funded under Title I of the Elementary and Secondary Education Act of 1965) directed toward providing compensatory reading education for those students who had failed to acquire the basic skills. Therefore, in terms of policy implications, perhaps the principal contribution of our paper is to provide a preliminary and necessarily tentative assessment of what the narrowly defined economic benefits would be of varying degrees of success in our nationwide efforts at compensatory reading education.

The chapter is organized as follows: In Section I we describe our basic models and data, and in Section II we present results from analyzing our data by way of a standard earnings function. In

---

*Mark Blaug, P.R.G. Layard, and George Psacharopoulos made valuable comments on an early version of this chapter, and the authors are particularly indebted to Henry M. Levin for helpful ideas and comments.

Sections III and IV we estimate the block recursive models we actually use to estimate benefits; Section III deals with the labor market part of the model and Section IV deals with the schooling-literacy part of the model. In Section V we develop our methodology for computing total benefits, and, using the empirical results of Sections III and IV, compute the benefits of schooling and compensatory education programs. Appendices provide more information about our data, and additional results based on some alternative approaches that are mentioned only briefly in the main text.

I. Models and Data

In this section we describe first the basic models we consider for analyzing our data, than describe the data itself. Since we have no information on a number of the variables that would, ideally, appear in an analysis of this sort, we conclude this section by discussing some of those missing variables and the implications their absence has for our analysis.

Models. Figure 1 presents schematically the alternative block recursive models we considered. The exogenous variables, those variables whose variations are not explained in the analysis -- sex, race, age, and parents' educations -- are assumed to determine schooling and literacy. We present, however, three alternative models for this process, which are labelled A, B, and C in Figure 1. Models A and B are strictly recursive, in that causation flows in a prespecified one-way direction; in Model A, schooling is assumed to determine literacy, and in Model B literacy determines schooling (on the assumption that continued success and willingness to stay in school is determined at least in part by reading competence). Model C is a simultaneous one, assuming literacy and schooling to be simultaneously determined. Model C thus includes A and B as special cases and, were it not for the problem of identification, our analysis would focus on Model C. However, for a variety of reasons, discussed more fully in the course of the paper, Model A seemed most suitable for analyzing our data set; we thus use Model A in the text of the paper; and its results are described in Section IV. In Appendix D, though, we present two-stage least squares estimates of Model C to be used for comparison with Model A.

The next major box to the right in Figure 1 schematizes the labor market model. Again there is the problem of whether to develop a simultaneous or recursive model, and again there are three alternatives, which we label 1, 2, and 3, with the obvious interpretations. In accord with most studies of labor supply based on survey data we end up by assuming wages to affect hours worked but not vice versa; our labor market results are thus based on Model 1, and our overall results on Model A-1. Estimation of Model 1 appears in Section III. We have again estimated the simultaneous model, Model 3, for comparison and the results of this estimation appear in Appendix C.

Since Models A and 1 are both recursive, the overall model we have chosen to estimate is strictly recursive, greatly simplifying problems of identification and estimation. We are mindful of the potential distortions this particular specification may have, and discuss its specific advantages and disadvantages at a number of points in the chapter. We also point out the direction of bias it could induce in estimating the relative benefits of schooling and compensatory reading education.

Variables and Data. In recent years there has been a considerable literature examining the effects of education on earnings. Most of these studies, however, have been based on samples that are inadequate in one or more of the following aspects: small sample size, too specialized a sample from which to form generalizations, or inadequate measures of education and ability. In general, most studies have been concerned with urban white males. Only in recent years have studies on the earnings of blacks and women begun to be undertaken. For example, studies of earnings of blacks have been made by Weiss (1970) and Welch (1973), and studies of earnings of women may be found in Kreps (1971), Hoffer (1973), Woodhall (1973), and Mincer and Polachek (1973).

The sample data used in the present study, known as the National Reading Performance Survey, were collected in 1973 for the U.S. Department of Health, Education, and Welfare through a contract to Educational Testing Service. There are several major advantages in using this sample in preference to the others. First, it is a national probability sample covering both men and women age 16 and over in all geographic areas. Second, in addition to the usual socioeconomic and other background variables which may determine earnings, data on educational level and reading competence are also available. The availability of reading competence data is, for reasons mentioned in the introductory paragraphs of the chapter, especially important. Although the effects of some measures of ability, such as IQ scores and Air Force Qualifying Test scores, on the level of earnings have been studied, it is plausible that reading competence is more subject to the influence of schools, and that the study of its effects are, therefore, of greater policy relevance.[1] Third, our sample data refer to 1972, more recent data than most of the data analyzed in recent literature, and thus deserve special attention. Finally, the sample contains information that allows us to estimate the labor market segment of the model of Figure 1.

The general survey design of the National Reading Performance Survey was based on a probability sampling model, using households as basic sampling units. Everyone in the selected household, 16 years of age and older, was to be interviewed. The necessity of callbacks and persistence to acheive high completion rates was stressed; this resulted in an overall response rate of approximately 70%. The survey instruments for each respondent consisted of a brief demographic questionnaire and one of ten books each containing 17 reading tasks. Of the 7,866 persons interviewed in the survey, 270 responded only to the demographic questionnaire because they were visually handicapped, unable to read the headlines in a newspaper, or simply refused to answer the reading tasks.[2]

The subsample used for the analysis reported in this paper is limited to individuals of age 25 to 60 who reported some earnings for 1972 and who are either white or black. Furthermore, individuals on whom information was incomplete were also eliminated from our subsample. As a result of introducing these restrictions, the actual subsample size used in this study becomes 2,308 individuals. The means and standard deviations of variables in our total sample and ~~in each of our four race-sex subsamples are presented in Table 5.1.~~ Appendix A contains the correlation matrices for the total sample and each of the subsamples.

As can be seen from the table, in our sample approximately 6% of the respondents are black and 41% are women. The mean age is about 39. The data also clearly show that men earned considerably more than women as a result of working somewhat longer hours at considerably higher wage rates. Years of schooling and reading scores are generally lower for blacks, especially black males. Variables whose meanings are not self-explanatory in Table 1 are discussed further below.

Y: Earnings for 1972 reported by the individuals interviewed, measured in thousand dollars;

Y_1: Wage rate computed by dividing earnings by work hours ($Y_1 = Y \div Y_2$);

Y_2: Work hours, measured in thousands of hours worked during 1972 (full-time workers who worked all year around are assumed to work 2,000 hours);

X_2: Reading scores, measured by standardized scores on one of ten sets of 17 reading tasks administered at the time of the survey (the items were all designed to measure basic literacy, and thus provide discrimination only among those with low reading competence);

X_3: Potential work experience, measured by subtracting schooling plus 5 from age ($X_3 = X_7 - X_1 - 5$);

X_6: Other income, all family incomes other than those earned by the individual interviewed, measured in thousands of dollars;

X_7: Age of individual interviewed;

X_{10}: Employment status of the individual interviewed (full-time salaried workers = 0, self-employed and part-time workers = 1).

Missing Variables. A number of variables that are plausibly important determinants of income were not available from the National Reading Performance Survey, and their omission raises cautions in interpreting our results. Four of the most important categories of missing variables are parental income, school quality, personality attributes of the respondent, and occupation of the respondent. Bowles and Nelson (1974) mention results from a study by Hauser, Lutterman, and Sewell (1971) that indicate parents' income to affect adult status independent of their education. Sewell and Hauser (1972) report direct effects of father's occupation on son's occupation, and father's income on son's. Thus our inclusion of only parents' education as a proxy for SES clearly limits our analysis.

A second category of variable missing from our data set is some measure of school input quality. Early work on the effects of school quality measures, e.g., Welch (1966), concluded that there were important effects; but Welch had available only highly aggregated data from which to draw this conclusion. More recent analyses using recursively structured earnings functions -- Ribich and Murphy (n.d.), Wachtel (1974) -- also find positive effects; much of the effect is through the influence of quality measures on years of educational attainment. Both the Ribich and Murphy and the Wachtel samples provide information only for males; their samples are further specialized in that Ribich and Murphy have data only for very recent entrants to the labor market (Project Talent data) and Wachtel only for high ability individuals (NBER/Thorndike-Hagen data). Nevertheless, the positive findings for these limited samples suggests the potential value of examining school quality measures in a probability sample such as ours.

An important school of thought -- perhaps best presented in Gintis (1971) -- maintains that the observed high correlation between schooling and earnings results not from the cognitive effects of schooling but rather from its effects on personality variables. Gintis persuasively makes the case that even after controlling for certain measures of cognitive outcome, schooling has a strong independent effect on earnings. (Our own results, even though we show literacy to have an important positive effect on the earnings of white males, are consistent with this conclusion of Gintis's.) However, to our knowledge, there exist no earnings functions that include personality measures as independent variables, and our data set allows no exception in this respect.[3] To the extent that affective outcomes do constitute an important fraction of the link between schooling and earnings, our analysis must be regarded as seriously incomplete.

A fourth shortcoming of our data set was lack of adequate information on the respondents' occupations. Occupation is an important intervening variable between background and schooling on the one hand, and wages and labor force participation on the other; the nature of this linkage has been of particular concern to sociologists, e.g., Duncan, Featherman, and Duncan (1972). The absence of data on occupation is of particular importance in our study because the differing returns to literacy by race and sex may, we hypothesize, result from interaction effects of literacy and occupation on income. We discuss this possibility in more detail later.

Thus there is a range of important questions that our analysis will be unable to address, and lack of information on some of these missing variables suggests caution in interpreting our results. Yet in spite of these weaknesses, our data set has a number of unique features that make its analysis worthwhile; most important of these are that it was generated from a probability sample of the U.S. population, it includes a measure of individual literacy, and it allows simultaneous study of educational attainment and labor force participation. We turn now to our results.

II. Empirical Earnings Function

Recent economic literature inquiring into the effects of education on earnings have generally followed the work of Schultz (1963), Becker (1964), and Mincer (1970). Although most of these studies have been concerned with the rate of return to education in the United States, similar studies have also been made for many other countries; for a review see Psacharopoulos (1973). While the primary concern of the present study is not to estimate an empirical earnings function, we believe it worthwhile to present our earnings function for comparison with the existing literature. There are several important questions that we shall attempt to answer in our analysis of empirical earnings functions. Among these are the questions of whether the earnings functions differ for blacks and whites, or for males and females; and if so, in what way and to what extent they differ. In addition, we shall consider some more specific questions such as whether the returns to education are different for blacks and whites, or for males and females. Finally, and central to our study, we shall also be interested in the effects of reading competence on the level of individual earnings, and these effects vary by race and sex.

Most empirical earnings functions in the literature have either employed a semi-logarithmic or simple linear function, using years of schooling, work experience, and the other socioeconomic variables as the explanatory variables. Occasionally, squared variables (or other transformations) are also included as explanatory variables. In general, assuming quadratic specifications, the empirical earnings function is generally specified as one of the following two convenient forms:

$$\text{Ln } Y = \alpha_0 + \sum_{i=1}^{k} \alpha_i X_i + \sum_{i=1}^{k} \alpha_{i+k} X_i^2 + U \quad (1)$$

or

$$Y = \beta_0 + \sum_{i=1}^{k} \beta_i X_i + \sum_{i=1}^{k} \beta_{i+k} X_i^2 + V \quad (2)$$

where α's and β's are the parameters to be estimated (some of them may be restricted to zero), k is the number of linear explanatory variables, and U and V are error terms, generally assumed to have zero mean and finite variance. Using these specifications, and applying ordinary least squares for estimation, we have obtained the empirical earnings function as reported in Tables 5.2 and 5.3 where the results of semi-logarithmic and simple linear specifications are reported separately.[4]

Psacharopoulas (1973) found in his survey of earnings functions no consistent empirical support for eq. (1) over eq. (2), and the linear results can be more easily understood or interpreted. However, strictly speaking, the choice between the log-linear specification (eq. 1) and the linear one (eq. 2) cannot be made simply by comparing goodness of statistical fit as represented by R^2's. Earlier, Mincer (1958, 1972) had advanced a theoretical argument for using eq. (1). More recently, Heckman and Polachek (1974), employing a Box and Cox's procedure, found eq. (1) to be empirically superior to eq. (2), using the 1960 and 1970 Census samples and the 1967 Survey of Economic Opportunity Data. For this reason, both empirical earnings function of eqs. (1) and (2) are reported.

According to both Tables 5.2 and 5.3 years of schooling is clearly a significant factor affecting earnings. The effects of reading scores on earnings are significantly positive only for white males. Apparently the effects of reading competence on earnings for blacks, and to some extent for white females, are unimportant. One plausible explanation for the differing effects of reading scores across subgroups is that there is a strong interaction between the effects of occupation and literacy on earnings. Literacy may be helpful in some occupations but not in others, and white males might have preponderant access to the occupations in which it is useful. As our sample contains only poor information on the respondents' occupation, we cannot test this hypothesis. Nonetheless, if it were correct, it would suggest that improving reading scores for, say, black females might still have potential economic benefits; realization of this potential would depend on their having better access to the appropriate occupations.

The effects of work experience on earnings are generally positive and have a general tendency to decline as the number of years and work experience increases. The magnitudes of these effects, however, can be estimated reliably only for whites, especially white males. Although we expect our measure of potential work experience may be less satisfactory for females whose work experience is more likely to exhibit a discontinuous pattern, we did not anticipate the considerable difference in the effects of work experience for white and black males that our results indicate.[5] Both the effects of father's and mother's education on earnings seem to be unimportant for all subsamples.[6] In some instances, the estimated effects of mother's education are negative, and in the case of black females even statistically significant. The effects of work hours on earnings are clearly statistically very significant, especially for whites as compared with blacks. The effects of race and sex on earnings are clearly important, judging from the fact that the dummy variables in the regressions computed from total sample are statistically significant and the regressions of different subsamples seem to be quite different.[7] Finally, since most studies of earnings functions are related to the empirical estimation of the rate of return, we may point out that rough estimates of the rates of return, to schooling, according to an approach suggested by Mincer, are provided by the regression coefficients associated with schooling variable in Table 5.2. Thus, the estimated rates of return to schooling are 7.7% and 8.7% for white males and females and 7.4% and 15.2% for black males and females respectively. Similar estimates may be derived from Table 5.3 by calculating $(\partial Y/\partial X_1)Y$ which also provides rough estimates of the rates of return to schooling.[8] According to this approach, the rates of return to schooling (at the mean earnings) are 7.4% and 9.7% for white males and females and 9.0% and 12.4% for black males and females.[9] We must point out, however, that these rough estimates fail to adjust for possible effects of schooling on other explanatory variables, such as reading scores and work experience. We shall consider this issue more fully later in our discussion of the economic benefits of schooling and compensatory reading.

III. Labor Market Analysis: Wage Rate and Work Hours

The results of empirical earnings functions shown in Tables 5.2 and 5.3, while they provide interesting information, can be difficult to interpret for some purposes. This is, in part because the effects of schooling and literacy on wage rate and work hours are intermingled in eqs. (1) and (2), and in part because of interdependence of schooling and literacy. In this section we analyze in more detail the structure of the labor market, and in the next section we deal with the interaction of schooling and literacy. To understand why the effects of schooling and literacy are intermingled, we must inquire into the meanings of the parameters α's and β's in eqs. (1) and (2). For simplicity, assuming eqs. (1) and (2) are strictly linear in explanatory variables, we can verify that

$$\alpha_i = \left(\frac{\partial \ln Y}{\partial X_i}\right) = \left(\frac{1}{Y_1}\right)\left(\frac{\partial Y_1}{\partial X_i}\right) + \left(\frac{1}{Y_2}\right)\left(\frac{\partial Y_2}{\partial X_i}\right)$$

$$= (Y/X_i) \rho_i (1 + \varepsilon)$$

$$\beta_i = \left(\frac{\partial Y}{\partial X_i}\right) = Y_2 \left(\frac{\partial Y_1}{\partial X_i}\right) + Y_1 \left(\frac{\partial Y_2}{\partial X_i}\right)$$

and $$= (1/X_i) \rho_i (1 + \varepsilon)$$

where $\rho_i = \left(\frac{X_i}{Y_1}\right)\left(\frac{\partial Y_1}{\partial X_i}\right)$ and $\varepsilon = \left(\frac{\partial Y_2}{\partial Y_1}\right)\left(\frac{Y_1}{Y_2}\right)$

are the elasticity of wage rate with respect to X_i and the elasticity of work hours with respect to wage rate (elasticity of labor supply?). Clearly the parameters α's and β's reflect not only the direct effect of exogenous variables on wage rate, but also their indirect effect on work hours through wage rate. In this section we shall, therefore, analyze the effects of various factors on wage rate and work hours by estimating the wage determination function and the work hours function separately. Ideally, an analysis of labor market should consider both demand and supply factors simultaneously, and one way of formulating such a model is to consider the wage determination function as the inverse demand function for labor and the work hours function as the supply function of labor.[10] Conceptually the wage determination function and the work hours function can, therefore, be regarded as a system of two simultaneous equations, where Y_1 and Y_2 are the two endogenous variables. In formal notation, using linear specifications, the wage determination function and the work hours function can be written as

$$Y_1 = \gamma_0 + \Sigma \gamma_i X_i + \gamma Y_2 + U \quad (3)$$

and

$$Y_2 = \delta_0 + \Sigma \delta_i X_i + \delta Y_1 + V \quad (4)$$

where γ's and δ's are the parameters to be estimated, and U and V are error terms.[11] For identification purpose, some of the parameters γ's and δ's must be restricted to be zero. The model as is formulated is a simultaneous model because neither $\gamma=0$ nor $\delta=0$ is necessarily imposed. One version of the simultaneous model has been estimated by both the Ordinary Least Squares (OLS) and the Two-Stage Least Squares (TSLS) procedures. The TSLS estimates, however, have been obtained only with an additional restriction in order to avoid a singular matrix in the second stage of computation. For this reason, in the following text only the results of a recursive model (assuming $\gamma=0$ and $\delta\neq0$) will be examined, and the results of a simultaneous model ($\gamma\neq0$ and $\delta\neq0$) are presented in Appendix C. Notice that in the simultaneous model we assume wage rate affects work hours and vice versa; in the recursive model we assume wage rate affects work hours but not vice versa.[12]

The empirical results of the wage determination function and the work hours function of the recursive model are presented in

Tables 5.4 and 5.5. Since the model is specified as recursive, OLS is an appropriate estimation procedure, and there is no need to employ TSLS or any other estimation procedure designed for estimating the parameters of a system of simultaneous equations.

The empirical results of Tables 5.4 and 5.5 show that R^2's of both the wage determination function and work hours function are considerably lower than those obtained for the earnings functions. Nevertheless, some of the individual coefficients are statistically highly significant, especially in the wage determination function and work hours function of white males. In general, schooling appears to be a significant factor in determining wage rates for all subsamples, and is also a significant factor in determining work hours of both white males and white females. The effects of reading scores on wage rate and work hours are generally insignificant, except for a negative effect on work hours for black females. There is some evidence indicating that reading scores probably have a slight effect on wage rate of white males and that their effects on work hours are positive for male workers but negative for female workers. The effects of work experience on wage rate and work hours appear to be more significant for whites than blacks, especially for white males. In general, wage rate appears to increase with work experience but at a smaller rate as experience increases, except for black females. Less experienced white males tend to work for longer hours than more experienced white males. The same is perhaps true for black females. A similar pattern, however, is not indicated for white females or black males.

The effects of father's education and mother's education on wage rate and work hours appear to be quite different for different subsamples. In general, the effects of father's and mother's education on wage rate are negative for white males, perhaps indicating a willingness to trade off income for status.[13] The effect of mother's education on wage rate is negative for blacks, especially for females. The effects of father's and mother's education on work hours are generally positive, though they are ususally not statistically significant except the effects of father's education for blacks. The effects of father's education on work hours for white females is negative and almost statistically significant, a result whose interpretation does not seem to be apparent. The effect of other income on work hours, which corresponds roughly with the effect of wife's income and husband's income for males and females respectively, is statistically significant only for whites. Nevertheless, the empirical results clearly indicate that such an effect is negative, as would be expected, for the work hours of female workers or the labor supply of married women. However, the effect of wage rate on work hours is negative for all subsamples, though only the coefficients for white males and females are statistically significant, implying a strong possibility of backward-bending labor supply curves. Finally, the effects of race and sex on wage rate and work hours are generally significant statistically, though the dummy variable for race is not significant in the work hours function computed from the total sample.

IV. Determinants of Educational Achievements: Years of Schooling and Reading Competence

In the previous section the effects of schooling and reading scores on both wage rate and work hours have been analyzed along with other background variables such as father's and mother's education as well as race and sex. The purpose of this section is to inquire further into the determinants of years of schooling and reading competence, both may be regarded as alternative measures of educational achievement. Conceptually the production function approach, which has been increasingly applied to educational processes, may be useful. However, because of the lack of school quality measures in our data, no attempt has been made to follow this approach in the following analysis. Our major concerns in this section are simply to determine what are the significant factors that may affect years of schooling and reading competence. In particular, we shall also be concerned with the questions of whether reading competence may be determined by years of schooling, and possibly, though perhaps unlikely in the present sample, vice versa. On the present sample, the measure of reading competence is obtained years after the respondents have left school.

In a general form, the educational achievement model may be specified as

$$X_1 = \lambda_o + \Sigma_{i=3}^{k} \lambda_i X_i + \lambda X_2 + U \qquad (5)$$

$$X_2 = \mu_o + \Sigma_{i=3}^{k} \mu_i X_i + \mu X_1 + V \qquad (6)$$

where X_1 and X_2 are years of schooling and reading scores, λ's and μ's are the parameters to be estimated, and i includes a given set of age and other background variables. The educational achievement model as specified above is a two-equation simultaneous model. As it stands, the schooling equation and the reading equation are not identified. Therefore, some additional restrictions on the parameters λ's and μ's must be imposed. The approach that has been employed in the present study for identification purpose is to restrict the parameters associated with father's education and mother's education in the schooling function to be the same, and also to restrict the parameters of father's education for the white and mother's education for the black in the reading equation to be zero. These restrictions are somewhat arbitrary and are based mainly on judgments derived from the preliminary empirical results. For this reason the empirical results of the simultaneous model of educational achievements will not be discussed here. They are presented, however, in Appendix D, because some of its results are interesting, despite the possible shortcomings of the identification procedure.

Rather than examining the empirical results of the simultaneous model (assuming $\lambda \neq 0$ and $\mu \neq 0$), the following discussion will be limited to those of the recursive model (assuming $\lambda = 0$ and $\mu \neq 0$) that implies schooling affects reading but not vice versa. This recursive model is not unreasonable, since our reading scores are measures of reading competence taken after individuals left their schools. Carnoy (1972) suggested another type of recursive model, which implies reading (or other measures of ability) affects schooling but not vice versa. This type of recursive model, as is supported by our preliminary empirical evidence, is less suitable for our sample.

The actual explanatory variables included in eqs. (5) and (6) are father's education, mother's education, age, race and sex.[14] Notice that reading scores are not included in eq. (5) but schooling is included in eq. (6) in the recursive model to be discussed. Because the selected educational achievement model is recursive, OLS can be applied to estimate the parameters of eqs. (5) and (6). It must be mentioned that we have treated the estimation of eqs. (5) and (6) separately from that of eqs. (3) and (4), partly because of our belief in the blockwise recursive nature of our specified models and partly because several difficulties were encountered in the simultaneous estimation of our labor market and educational achievement models, as was explained earlier. In any event, since the selected labor market and educational achievement models are both recursive, the OLS procedure can be appropriately applied to estimate the parameters of each equation separately.

The empirical results of the recursive educational achievement model (assuming $\lambda = 0$) are presented in Tables 5.6 and 5.7 for the determinants of schooling and reading scores respectively. In general, goodness-of-fits as represented by R^2's are reasonable. Most of the individual coefficients are statistically significant. The effect of age on schooling is statistically significantly negative for males of both black and white, with numerical magnitude for black males considerably larger than that of white males (in absolute values), reflecting the fact that the average increase of educational level perhaps has been the fastest for black males. The effect of age on reading scores, however, indicates a somewhat different interpretation: while younger persons seem to read better than the older persons among the white, the same does not appear to be true for the black. The evidence is perhaps the strongest for white females and the weakest for black females. While both father's education and mother's education appear to have a positive effect on schooling, their effects on reading scores seem to be quite different for whites and blacks.[15] The effects of father's and mother's education on reading are negative and positive respectively for whites, but the reverse is true for blacks. For whites it is the effect of mother's education on reading that is significantly positive, but for blacks it is the effect of father's education that is significantly positive, but for blacks it is the effect of father's education that is significantly positive. The effect of schooling on reading is statistically significant for all subsamples, but the magnitude of the effect is almost twice as large for blacks as for whites. Finally, the effects of race and sex on schooling and reading appear to be important. This is apparent from the statistical significance of the dummy variables in the regressions based on total sample and from the differences among the regressions based on different subsamples.

V. Economic Benefits of Schooling and Compensatory Reading

The purpose of this section is to demonstrate how the framework of our previous labor market and educational achievement models can be combined to explain the sources of earnings difference, and thus how the total economic benefits of schooling and compensatory reading may be assessed.[17]

Most recent studies of the rate of return to education attempted to estimate an empirical earnings function using a semi-logarithmic form or a simple linear form similar to eq. (1) or (2) respectively. In an early study, Mincer (1958) suggested that the rate of return to education can be estimated by the coefficient of years of schooling in a semi-logarithmic form, and most existing studies seem to show that such an approach can indeed be useful. More recently, however, some of the possible limitations of such an approach have become apparent. For example, the estimated rate of return to schooling may be biased because of missing variables that are likely to be correlated with schooling. Griliches and Mason (1972) have examined this problem by considering the effect of ability and found the bias to be insignificant. However, there are some other problems. As more variables are included in the earnings function, it also becomes apparent that a reasonable estimate of the rate of return to schooling cannot be obtained without explicitly taking into account the interrelationships among the explanatory variables. Welch (1973), in his recent study of black-white difference in returns to schooling has attempted to deal with this problem by considering a set of auxiliary regressions that explain the interrelationships among the explanatory variables. In addition, the questions regarding whether the dependent variable should be earnings or wage rate and whether work hours should be an explanatory variable in an earnings function have been raised. These and other questions suggest that a study of earnings can perhaps be more meaningful handled by a more detailed labor market analysis, using a multiple-equation approach. In his recent study of wage rate and work hours, Hall (1973) demonstrated the potential of this approach, though his major concern was not directly related to the earnings function. Several other studies, e.g., Weiss (1970) and Blinder (1973), also recognized the importance of a multiple-equation approach. None of the existing studies, however, have attempted to separate the effects of schooling or other determinants of earnings into the effects due to changes in wage rate and work hours.

Although the economic benefits of schooling and reading can be estimated directly from the empirical earnings functions such as eqs. (1) and (2), probably with some adjustments as was done by Welch (1973), such an approach will not be followed because, as was pointed out previously, it does not provide a framework for identifying whether wage rate or work hours may be the main source of difference in earnings. Therefore, instead of simply relying on our empirical earnings function, we shall use the definition of earnings that is the product of wage rate and work hours ($Y = Y_1 Y_2$) combined with the empirical results of our labor market and educational achievement models for subsequent discussion. Formally, our analytical structure consists of the definition of earnings, the wage determination and work hours functions, i.e., eqs. (3) and (4), and the determinants of schooling and reading, i.e., eqs. (5) and (6).

To facilitate discussion, we first discuss how the partial benefits of schooling and reading or any other factor affecting either wage rate or work hours may be evaluated, ignoring the interrelationships among the determining factors, such as those examined in the educational achievement model. Later we shall consider how these partial benefits may be combined in order to obtain the full economic benefits of schooling and reading, using the empirical relationships of the educational achievement model. It can be verified that the reduced form of the labor market model represented by eqs. (3) and (4) is

$$Y_1 = \{ 1/(1 - \alpha\beta)\}\{(\alpha_0 + \alpha\beta_0) + \Sigma_{i=1}^{k} (\alpha_i + \alpha\beta_i) X_i + \Sigma_{i=1}^{k} (\alpha_{i+k} + \alpha\beta_{i+k}) X_i^2 \}, \tag{7}$$

$$Y_2 = \{ 1/(1 - \alpha\beta)\}\{(\beta_0 + \beta\alpha_0) + \Sigma_{i=1}^{k} (\beta_i + \beta\alpha_i) X_i + \Sigma_{i=1}^{k} (\beta_{i+k} + \beta\alpha_{i+k}) X_i^2 \}, \tag{8}$$

which are obtained simply by eliminating Y_2 from eq. (3) and Y_1 from eq. (4). Thus, from the definition of earnings, $Y = Y_1 Y_2$, the partial benefits of any determinant of earnings can be written as

$$\left(\frac{\partial Y}{\partial X_i}\right) = Y_2 \left(\frac{\partial Y_1}{\partial X_i}\right) + Y_1 \left(\frac{\partial Y_2}{\partial X_i}\right), \quad i = 1,2,\ldots k \tag{9}$$

where

$$\left(\frac{\partial Y_1}{\partial X_i}\right) = \alpha_i + \alpha\beta_i + 2\alpha\beta_{i+k} X_i$$

and

$$\left(\frac{\partial Y_2}{\partial X_i}\right) = \beta_i + \beta\alpha_i + 2\beta\alpha_{i+k} X_i$$

according to eqs. (7) and (8). In our empirical results presented in Tables 5.3 and 5.4, most of the parameters α_{i+k} and β_{i+k} are restricted to be zero, since the only squared variable is work experience. The partial benefits defined above can be clearly decomposed into two components representing wage-rate effect and work-hours effect respectively. Notice that these partial effects depend on the specific forms of the wage determination and work hours functions. Moreover, they depend on which of the other determinants are held constant.

In the discussion of partial benefits we have treated schooling, reading, and work experience in the same way as we have treated other exogenous variables such as father's education and mother's education, which are clearly exogenous and beyond the choice of the individuals whose earnings are being analyzed. To evaluate more fully the economic benefits of schooling and reading, it is necessary for us to take into account some possible interrelationships among the explanatory variables that so far have been treated as exogenous. In general, the full benefit of any determinant of earnings can be defined as

$$\left(\frac{dY}{dX_i}\right) = \Sigma_{j=1}^{k} \left(\frac{dX_i}{dX_j}\right)\left(\frac{\partial Y}{\partial X_i}\right), \; i=1, 2, \ldots k \tag{10}$$

where $\left(\frac{\partial Y}{\partial X_i}\right)$ are partial benefits previously defined in eq. (9). Thus full benefits are simply weighted sums of partial benefits, dX_i/dX_j being the weights.

Among the many possible interrelationships among the explanatory variables, the most important ones are the definition of work experience and the interrelationships studied in our educational achievement model represented by eqs. (5) and (6). Assuming these are the only interrelationships among the explanatory variables, specific measures of full benefits of schooling and reading can then be computed. From the definition of work experience as a function of age, i.e., $X_3 = X_7 - (X_1 + 5)$, we have $dX_3/dX_1 = -1$. From the empirical functions of the determinants of schooling and reading we know that $dX_1/dX_2 = \lambda$ and $dX_2/dX_1 = \mu$, according to eqs. (5) and (6) respectively. Therefore, the full benefits of schooling and reading competence measured in terms of incremental annual earnings may be defined explicitly as

$$\left(\frac{dY}{dX_1}\right) = \left\{\left(\frac{\partial Y}{\partial X_1}\right) - \left(\frac{\partial Y}{\partial X_3}\right)\right\} + \mu\left(\frac{\partial Y}{\partial X_2}\right), \tag{11}$$

and

$$\left(\frac{dY}{dX_2}\right) = \lambda\left\{\left(\frac{\partial Y}{\partial X_1}\right) - \left(\frac{\partial Y}{\partial X_3}\right)\right\} + \left(\frac{\partial Y}{\partial X_2}\right), \tag{12}$$

where in both cases the first term represents the benefits attributable to schooling and the second term to reading. Therefore, the first term in eq. (11) and the second term in eq. (12) may be regarded as direct benefits of schooling and reading competence, and the second term in eq. (11) and the first term in eq. (12) their corresponding indirect benefits. Notice that, in general, these benefits are functions of age, schooling, and other background variables, since partial benefits are functions of wage rate and work hours.

Full benefits of schooling and reading competence as defined in eqs. (11) and (12) can be expanded and rearranged as

$$\left(\frac{dY}{dX_1}\right) = Y_2\left\{\left(\frac{\partial Y_1}{\partial X_1}\right) - \left(\frac{\partial Y_1}{\partial X_3}\right) + \mu\cdot\left(\frac{\partial Y_1}{\partial X_2}\right)\right\} + Y_1\left\{\left(\frac{\partial Y_2}{\partial X_1}\right) - \left(\frac{\partial Y_2}{\partial X_3}\right) + \mu\left(\frac{\partial Y_2}{\partial X_2}\right)\right\}, \tag{13}$$

$$\left(\frac{dY}{dX_2}\right) = Y_2\left\{\lambda\left[\left(\frac{\partial Y_1}{\partial X_1}\right) - \left(\frac{\partial Y_1}{\partial X_3}\right)\right] + \left(\frac{\partial Y_1}{\partial X_2}\right)\right\} + Y_1\left\{\lambda\left[\left(\frac{\partial Y_2}{\partial X_1}\right) - \left(\frac{\partial Y_2}{\partial X_3}\right)\right] + \left(\frac{\partial Y_2}{\partial X_2}\right)\right\}, \tag{14}$$

where the first terms are the effects of wage rate on full benefits of schooling and reading, and the second terms the corresponding effects of work hours. The decompositions of full benefits of schooling and reading into wage-rate and work-hours effects as suggested in eqs. (13) and (14) are not only useful in themselves but also convenient in order to compute some other measures of benefit. For example, similar to Eckaus (1973), alternative measures of benefits may be computed by assuming work-hours are fixed at the same level for all individuals. Thus, adjusted benefits of schooling and reading competence may be computed by dividing the first terms of eqs. (13) and (14) by the ratio of observed work hours to the fixed, say 2,000 hours, and ignoring the second terms representing the effects of work hours. An implicit assumption used in these measures is that individuals always work full time, either in the labor market, as self-employed, or in household production. These measures of adjusted benefits, though they clearly have some limitation, may be useful especially in indicating maximum benefits of schooling or reading.

So far we have discussed benefits of schooling and reading only in terms of incremental annual earnings. We have pointed out these benefits are in general functions of age, schooling, and other background variables. By holding all other variables constant, for example, at the observed mean levels and letting only age vary, we can construct a stream of annual earnings increments realizable at any given age due to an incremental change in schooling or reading. The present value of this stream of benefits, discounted at some appropriate discount rate, provides a more complete measure of the economic benefits of schooling or reading. Formally the present values of full benefits of schooling and reading may be computed from

$$\int_{n}^{m} \left(\frac{dY}{dX_1}\right) e^{-r(X_7-n)}\, dX_7 \tag{15}$$

and

$$\int_{n}^{m} \left(\frac{dY}{dX_2}\right) e^{-r(X_7-n)}\, dX_7 \tag{16}$$

where n is the current age and m is the retirement age. For practical purpose, the present values of full benefits of schooling and reading may be computed by discrete approximations of eqs. (15) and (16), so that indefinite integration may be avoided.

Finally, it must be pointed out that benefits of schooling and reading previously discussed are all marginal benefits reflecting changes in benefits due to changes in one unit of schooling or reading (evaluated at a given schooling or reading level). The total benefits of schooling or reading, measured in terms of annual earnings, due to changes over several units of schooling or reading can also be computed as

$$\int_{n}^{m} \left(\frac{dY}{dX_1}\right) dX_1 \tag{17}$$

and

$$\int_{n}^{m} \left(\frac{dY}{dX_2}\right) dX_2 \tag{18}$$

where n and m are the initial schooling or reading level and the targeted schooling or reading level respectively. Clearly corresponding average benefits are obtained simply by dividing eqs. (17) and (18) by m-n, representing the range of change in schooling or reading. These measures are particularly useful to answer such questions as: what are the possible economic benefits of increasing the level of schooling from n years to m years or the level of reading competence from n to m standardized score? Given the distribution of the initial schooling or reading competence, the possible benefits of a given educational program that would raise the schooling or reading level of all individuals to a given targeted schooling or reading level can also be computed. For example, we can compute the economic benefit of a compensatory reading program that would raise the reading competence of all individuals whose scores are under a given targeted level, say, the present national mean (zero in standardized scores). Although we have discussed the concept of average and total benefits only in terms of annual earnings, these same concepts can be applied to the present values explained in eqs. (15) and (16). That the numerical value of such a computation should be used with extreme caution goes without saying. One is both extrapolating from marginal to large changes and ignoring the possibility of non-optimizing producer choice or market signaling effects (Spence, 1974, Chapters 3 and 4). Nonetheless, we feel the computations do place a rough upper limit on the total benefits to be expected. The empirical results of the benefits of schooling and reading based on the concepts discussed above are presented in Tables 5.8 and 5.9. The results presented here are based on the empirical relationships given in Tables 5.4 through 5.7 assuming recursive structures for both labor market and educational achievement models. It should be noted that, since our assumed recursive structure has schooling affecting reading, but not vice versa, there may be some tendency to overstate the relative benefits of schooling and understate those of compensatory reading education.

In Table 5.8, the estimates of alternative measures of private benefits of schooling are presented. The table is divided into two parts: the first part gives the estimates of alternative measures of marginal benefit for an additional year of schooling at approximately the high school level, and the second part provides the estimates of total benefits for a representative individual and the nation as a whole for two hypothetical compulsory educational programs. The estimates of partial and full benefits of schooling, as defined in eqs. (9) and (11), are computed at the mean levels of all explanatory variables. In general, these two measures are very close to each other, with full benefits somewhat lower than partial benefits largely because of adjustments for foregone benefit due to experience. The estimated full benefits, measured in terms of increases in annual earnings, are the highest for white males ($1,121) and the lowest for white females ($517). The corresponding full benefits for black males and females are $663 and $783 respectively. When the full benefits are decomposed according to eq. (13) into the wage-rate and work-hours effects, it is apparent that most effects are due to wage rate rather than work hours. It is interesting to note that most estimated work-hour effects are negative, except for white females. The estimated adjusted benefits, similar to Eckaus (1973), are computed by dividing the first terms of eq. (13) by the ratio of actual work hours to full-time work hours (2,000 hours). Because the ratios of actual work hours to full-time work hours are close to one, and the work-hour effects are generally small, the results of adjusted benefits are not very different from the corresponding original estimates of full benefits. The present values of full benefit streams are computed at age 18, assuming retirement at age 65. The effect of discount rate on the magnitude of present value is shown by providing results for zero, 5% and 10% discount rates. It is important to note that the present values with 10% discount rate are perhaps very close to the mean earnings of individuals at age 18, suggesting that the private rates of return to schooling are approximately 10%, except for black females whose rate of return to schooling appears to be somewhat higher.

The estimated total benefits of two hypothetical compulsory educational programs must be received with great caution. The estimated benefits are based on the assumption that all individuals (age 25 and over) whose educational levels are lower than the targeted level (either high school or college graduation) were able to complete the compulsory education at the targeted level, and that they were able to obtain the same earnings as presently observed for the targeted levels of schooling. It is also assumed that no benefit (or loss) will occur to individuals whose level of schooling are already above the targeted level. The total benefits for a representative individual of the hypothetical compulsory educational programs are computed by

$$\int_{o}^{m} \left(\frac{dY}{dX_1}\right) f(X_1 | X_1 < m) dX_1 \qquad (19)$$

where $f(X_1|X_1<m)$ is the conditional distribution of individuals by educational level, and m is either 12 or 16, corresponding to high school and college graduation respectively. These total benefits are in effect measures of average benefits of all individuals whose education is under the given targeted level. As the table shows, the estimated benefits of the hypothetical compulsory high school education for a representative individual are the highest for white males ($3,810) and the lowest for white females ($1,370). The corresponding estimated benefits for black males and females are $2,580 and $1,940 respectively. The estimated benefits of the hypothetical compulsory college education for a representative individual can be interpreted analogously. Finally, the national program benefits are computed simply by multiplying the representative individual benefits by the corresponding total numbers of individuals completing less than high school or college education. The actual figures used for the numbers of individuals (age 25 and over) completing less than high school or college education are for 1970 taken from the <u>Statistical Abstract of the United States 1972</u>, No. 168. As the table shows, the estimated national program benefits of the hypothetical compulsory high school education are substantial: approximately \$76 billion and \$40 billion for white males and females and \$8 billion and \$7 billion for black males and females respectively. The estimated national program benefits of the hypothetical compulsory college education are even higher. It must be recognized, however, that practically such hypothetical programs cannot be realistically implemented.

We have so far discussed only our estimates of benefits of schooling as shown in Table 5.8. The corresponding estimates of benefits of reading competence are reported in Table 5.9, which is also divided into two parts: marginal benefit and total benefit. The estimates of partial and full benefits, as defined in eqs. (9) and (12), are also computed at the mean levels of all explanatory variables. The partial and full benefits, shown under the heading of marginal benefit, are identical because the underlying educational achievement model is recursive, i.e., $\lambda=0$ in eq. (12). As the table shows, the benefits of reading competence for males are larger than for females. In fact, our result shows that the benefits of reading competence is negative for black females. Whether this result can be taken seriously is, however, not clear to us. It is important to note that when full benefits are decomposed according to eq. (14) into wage-rate and work-hours effects, we find that the wage-rate effect is more important for white males but the work-hours effect is more important for black males. In addition, it is interesting to note that the wage-rate and work-hours effects are in opposite direction and almost cancelling the effects of each other completely for white females. The interpretation of the estimated adjusted benefits and present values at various discount rates are analogous to those of Table 5.8.

In the second part of Table 5.9, the estimated total benefits of two hypothetical compensatory reading programs, with low and high reading targets, are provided. We must stress that these estimates, like the similar estimates for the two hypothetical compulsory educational programs given in Table 5.8, must be received with great caution. The total benefits for a representative individual of the hypothetical compensatory reading programs are computed by

$$\int_{-\infty}^{m}\left(\frac{dY}{dX_2}\right) f(X_2|X_2<m)\, dX_2 \tag{20}$$

where $f(X_2|X_2<m)$ is the conditional distribution of individuals by standardized reading score, and m is either -1 or zero, corresponding to a low or high target compensatory reading program. As the table shows, the estimated benefits of the low target compensatory reading

program for a representative individual are \$323 and \$273 for white males and black males respectively, and only \$12 for white females and negative for black females. The estimated benefits of the high target compensatory reading program for a representative individual are somewhat higher as may be expected, except for black females.

The national program benefits are computed by multiplying the representative individual benefits by the corresponding estimated numbers of individuals whose reading competence, measured by standardized scores, are likely to fall below -1 or zero, using the frequency distributions of the National Reading Performance Survey and population figures (age 25 and over) for 1970 obtained from Statistical Abstract of the United States 1972, No. 168. The total benefits of a national program which raises everyone's reading competence to a level represented by -1 of standardized reading score are estimated to be approximately \$2 billion and \$71 million respectively for white males and females and approximately \$609 million and negative respectively for black males and females. The estimated benefits of the high reading target national program are, as expected, generally higher. Their interpretations are analogous, and, therefore, need no further explanation. Finally, we wish to emphasize again that these estimated benefits are very tentative. They may be biased downward for one reason, but biased upward for another reason.[18] Furthermore, like the hypothetical compulsory educational programs discussed previously, the goals of the hypothetical compensatory reading programs may be practically infeasible to achieve.

FOOTNOTES

1. We are aware of only a few prior studies dealing with the relation between literacy and earning. One is by Carnoy and Lockheed-Katz (1971) using Brazilian data; while they had insufficient information to specify an earnings function, they did find a positive association between literacy and earnings.

2. See R. Murphy (1973) for a detailed discussion of how the test instruments were developed and implemented.

3. One series of studies of correlations between personality variables and income, though restricted to graduates of the Master of Business Administration (MBA) program of the Stanford Graduate School of Business, does provide direct support for the Gintis position. Harrell (1969, 1970) and Harrell and Harrell (1974) found that high earnings MBSs tend to have more "ascendant" personalities and were "...overwhelmingly in the socially desirable direction on the personality measures" (Harrell, 1969, p. 461). Harrell and Harrell found a significant negative (simple) correlation between verbal score and earnings of MBAs, and attributed this to differences they found in personality. It would be of interest to ascertain the extent to which this finding would hold up in a multivariate (i.e., earnings function) analysis.

4. Although experience-squared is included as an explanatory variable in our estimations, eq. (2) is referred to as simple linear for convenience. In our early analyses, we have estimated the earnings function with cross-product terms attempting to explain the interactions between schooling and reading and between experience and reading. According to the results of semi-log earnings function, most of these interactions are positive. While the experience and reading interaction is more significant for white males, the schooling and reading interaction appears to be more significant for white females as well as for blacks of both sexes.

5. The term "work experience," throughout the present study, should be understood as potential work experience as previously defined. For males, potential and actual work experience are probably very close; females, however, after marriage, spend less than half their lifetime in the labor market on the average, and our data are unable to provide information on the actual amount. Mincer and Polachek (1974) have used the National Longitudinal Survey data to estimate the effects of actual labor market participation on women's earnings.

6. This result is in general consistent with the evidence provided by existing studies which indicates that the most important parental influences on the adult earnings of their children are indirect rather than direct. Thus, the effects of parents' education on schooling, and, to a lesser extent, literacy does make a contribution to future earnings, but the effects can be traced only through estimation of the type of recursive system estimated later in this paper. For a review of some of the existing literature, see C.R. Hill and F.P. Stafford (1974).

7. More rigorous testing procedure along the line suggested by G. Chow (1960) has not been performed. Our maintained hypothesis is that each subsample should be represented by an earnings function of its own.

8. This is because the rate of return to schooling, according to eq. (1) suggested by Mincer (1958), is $\partial \ln Y/\partial X_1$ which equals $(\partial Y/\partial X_1)/Y$.

9. Our results are thus consistent with those of Welch (1973) - that returns to education are now as high for blacks as whites. This is in contrast to earlier findings, but the more recent data used by Welch and by us suggests that there has been a change over time.

10. This simply reflects a particular normalization rule. The idea has been indicated by R. Hall (1973). No satisfactory empirical result, however, has been provided.

11. Although the same notation is used for the error terms of eqs. (1) and (3), and similarly for eqs. (2) and (4), they are in general different.

12. These problems of identification are standard in the labor supply literature that uses survey data; see, for discussion, the papers in the volume edited by Cain and Watts (1973) or Metcalf, Nickell, and Richardson (1974).

13. Henry Levin suggested the potential importance of this tradeoff to us.

14. In preliminary analyses we constructed, as an alternative to mother's and father's education, a measure of their education <u>relative</u> to what the education of a person their age would be, using a prediction of their age based on the respondent's age. This transformation affected the results in no substantial or consistent manner, so we returned to the more simple education variable.

15. Recall that the reading scores were constructed as "standardized scores," since ten different test booklets were administered. More specifically, the reading scores are defined as $(P_{ij} - \bar{P}_j)/S_j$, where P_{ij} is the proportion of right answers for i individual using j booklet and $\bar{P}_j$ and S_j are mean and standard deviation of P_{ij}. These standardized scores were based on all items in each booklet. Our supplementary study on the possible effect of deleting some "inappropriate" items on the results of our analysis indicates that such an effect can be expected to be relatively minor, since the correlations of standardized scores based on all items and "selected items" only are highly correlated. Transformations of reading scores were also experimented with. In particular, a transformation of reading score was defined as $-N_{ij}\ (1 + k)^{N_{ij}}$ where N_{ij} is the number of wrong answers for the i individual using j booklet divided by the ratio of the mean of the number of wrong answers for j booklet to that of all booklets, and k is a given constant, which was assigned a value ranging from -0.5 to 0.5. The results corresponding to eqs. (5) and (6) generally suggest that there is no significant difference among alternative transformations of reading scores. In addition, in Appendix B, we also examine the effects of substituting discontinuous variables for schooling and reading scores.

16. For the results of some other studies on the effect of parent's education on schooling, see C.R. Hill and F.P. Stafford (1974) or Woodhall (1973, p. 288).

17. The analytical framework developed here may have other potentially significant applications, for example, in analyzing an important issue on the sources of inflation and real economic growth. The analogy between this problem and the one discussed in the text is apparent, since inflation and real output correspond to wage and work hours respectively.

18. For example, these estimates of total benefits of reading competence may be biased downward because of our imposed assumption of "linear" effect, a possibility examined more fully in our Appendix B. On the other hand, these estimates may be biased upward because of our failure to consider the issue in a general equilibrium framework.

REFERENCES

Becker, G.S. Human Capital: A Theoretical and Empirical Analysis. New York: National Bureau of Economic Research, 1964.

Blinder, A.S. "Wage Discrimination: Reduced Form and Structural Estimates." Journal of Human Resources, 1973, 8, 456-471.

Bowles, S., and Nelson, V.I. "The 'Inheritance of IQ' and the Intergenerational Reproduction of Economic Inequality." Review of Economics and Statistics, 1974, 56, 39-51.

Cain, G.G., and Watts, H.W. (Eds.), Income Maintenance and Labor Supply. Chicago: Rand McNally, 1973.

Carnoy, M. "The Social Benefits of Better Schooling." Occasional Papers in the Economics and Politics of Education. Stanford University. March, 1972.

Carnoy, M., and Lockheed, M. "Explaining Differentials in Earnings Among Large Brazilian Cities." Urban Studies, 1971, 8, 21-37.

Duncan, O.D., Featherman, D.L., and Duncan, B. Socioeconomic Background and Achievement. New York and London: Seminar Press, 1972.

Eckaus, R.S. "Estimation of the Returns to Education with Hourly Standardized Income." Quarterly Journal of Economics, 1973, 121-131.

Gintis, H. "Education Technology and the Characteristics of Worker Productivity." American Economic Review, 1971, 61, 266-279.

Griliches, Z., and Mason, W.M. "Education, Income and Ability." Journal of Political Economy, 1972, 80, 74-103.

Hall, R.E. "Wages, Income, and Hours of Work in the U.S. Labor Force." In Income Maintenance and Labor Supply, edited by G.G. Cain and H.W. Watts. Chicago: Rand McNally, 1973.

Harrell, T.W. "The Personality of High Earning MBA's in Big Business." Personal Psychology, 1969, 22, 457-463.

Harrell, T.W. "The Personality of High Earning MBA's in Small Business." Personal Psychology, 1970, 23, 369-375.

Harrell, T.W., and Harrell, M.S. "The Relation of Verbal and Quantitative Scores to Manager Success." Unpublished manuscript. August, 1974.

Hauser, R., Lutterman, K., and Sewell, W. "Socioeconomic Background and the Earnings of High School Graduates." Unpublished manuscript. August, 1971.

Hill, C. R., and Stafford, F. P. "Family Background and Lifetime Earnings." Unpublished manuscript. May, 1974.

Hoffer, S. N. "Private Rates of Return to Higher Education." Review of Economics and Statistics, 1973, 55, 482-486.

Kreps, J. Sex in the Marketplace: American Women at Work. Baltimore and London: Johns Hopkins Press, 1971.

Metcalf, D., Nickell, S., and Richardson, R. "The Structure of Hours and Earnings in British Manufacturing Industry." London School of Economics. Unpublished manuscript. February, 1974.

Mincer, J. "Investment in Human Capital and Personal Income Distribution." Journal of Political Economy. August, 1958.

Mincer, J. "The Distribution of Labor Income: A Survey With Special Reference to the Human Capital Approach." Journal of Economic Literature, 1970, 53, 1-26.

Mincer, J. Schooling, Experience and Earnings. New York: National Bureau of Economic Research, 1972 (mimeographed).

Mincer, J., and Polachek, S. "Family Investment in Human Capital: Earnings of Women." Paper presented at Population Conference, Chicago, 1973.

Murphy, R. T. Adult Functional Reading Study. Final Report. Washington, D. C.: U. S. Department of Health, Education, and Welfare; National Institute of Education, 1973.

Psacharopoulos, G. Return to Education: An International Comparison. San Francisco and Amsterdam: Jossey-Bass/Elsevier, 1973.

Ribich, T., and Murphy, J. "The Economic Return to Increased Educational Spending." Unpublished manuscript (n.d.).

Schultz, T. W. The Economic Value of Education. New York: Columbia University Press, 1963.

Sewell, W. H., and Hauser, R. M. "Causes and Consequences of Higher Education: Models of the Status Attainment Process." American Journal of Agricultural Economics, 1972, 54, 851-861.

Spence, A. M. Market Signaling. Cambridge, Mass.: Harvard University Press, 1974.

Wachtel, P. "The Return to Investments in Higher Education: Another View." In Education, Income, and Human Behavior, edited by F. T. Juster. Carnegie Commission of Higher Education and National Bureau of Economic Research, 1974.

Weiss, R. D. "The Effect of Education on the Earnings of Blacks and Whites." Review of Economics and Statistics, 1970, 52, 150-159.

Welch, F. "Measurement of the Quality of Schooling." American Economic Review, 1966, 56, 379-392.

Welch, F. "Black-White Differences in Returns to Schooling." American Economic Review, 1973, 63, 893-907.

Woodhall, M. "The Economic Returns to Investment in Women's Education." Higher Education, 1973, 2, 275-300.

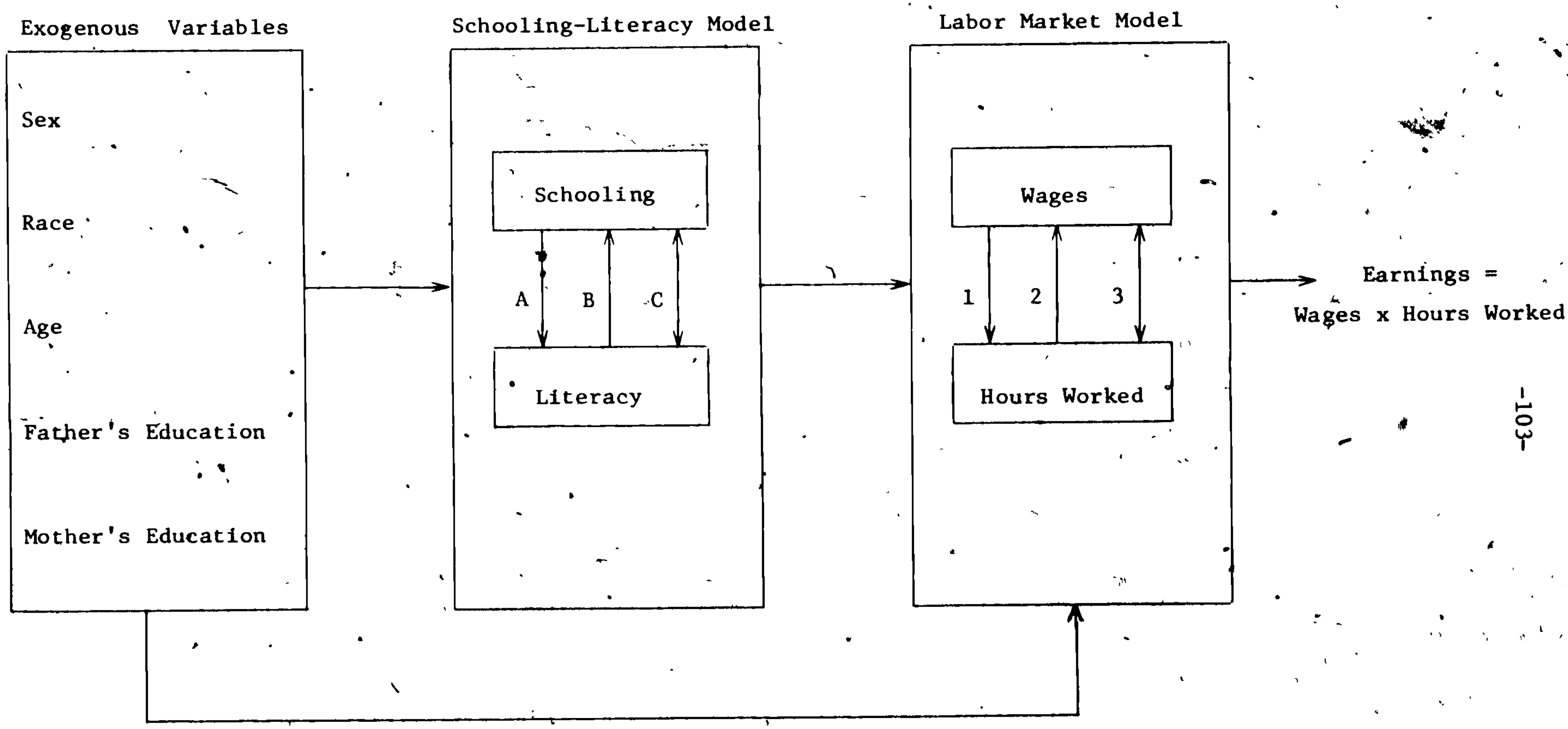


Figure 1

Block Recursive Models of Determinants
and Effects of Schooling and Literacy

TABLE 5.1

SAMPLE MEANS AND STANDARD DEVIATIONS OF VARIOUS SUBSAMPLES IN NATIONAL READING SURVEY, 1972

(AGE 25 to 60, WHITE AND BLACK)

	Variable	Total	White Male	White Female	Black Male	Black Female
Y	Earnings (thousands of dollars per year)	9.5973 (6.2233)	12.5931 (6.0149)	5.7527 (4.1263)	7.6585 (4.1986)	5.2938 (2.7634)
LnY	Earnings (log)	1.9967 (0.8155)	2.3982 (0.5655)	1.4660 (0.8160)	1.8660 (0.6306)	1.4862 (0.6688)
Y_1	Wage Rate (dollars per hour)	5.4595 (4.8225)	6.7177 (5.0562)	3.9292 (4.2174)	4.1087 (2.0859)	3.0780 (1.5818)
Y_2	Work Hours (thousands of hours per year)	1.7953 (0.4614)	1.9297 (0.2674)	1.6003 (0.5978)	1.8481 (0.3729)	1.7577 (0.4699)
X_1	Schooling (years)	12.7943 (2.8502)	12.9966 (3.0154)	12.6636 (2.5217)	11.1081 (3.1610)	12.5001 (2.5892)
X_2	Reading (standardized score)	-0.0001 (1.0003)	0.0627 (0.9655)	0.0580 (0.8922)	-1.1459 (1.4568)	-0.7685 (1.3172)
X_3	Experience (years of potential job experience)	21.5528 (11.1516)	21.0706 (11.0995)	22.2741 (11.1426)	21.6378 (12.0450)	21.0488 (10.6809)
X_3^2	Experience-squared	588.8828 (526.4469)	567.1686 (522.6717)	620.2948 (521.6531)	613.2794 (627.0424)	557.1324 (517.4029)
X_4	Father's Education (years)	9.0333 (4.2195)	8.9992 (4.3023)	9.2169 (4.1228)	7.8690 (3.4863)	8.4834 (4.3956)
X_5	Mother's Education (years)	9.3382 (3.7903)	9.3932 (3.8704)	9.3296 (3.7008)	8.4974 (3.4604)	9.3252 (3.6612)
X_6	Other Income (thousands of dollars per year)	8.1233 (12.0290)	6.2262 (11.4253)	11.3307 (12.5611)	3.2470 (6.7874)	7.0111 (10.7111)
X_7	Age (years)	39.3472 (10.2675)	39.0672 (10.1515)	39.9378 (10.3722)	37.7460 (10.4891)	38.5488 (10.3696)
X_8	Race (black = 0, white = 1)	0.0589 (0.2355)	---	---	---	---
X_9	Sex (female = 0, male = 1)	0.4131 (0.4924)	---	---	---	---
X_{10}	Employment Status (part-time=0, full time=1)	0.2998 (0.4582)	0.2168 (0.4121)	0.4315 (0.4953)	0.1744 (0.3795)	0.2709 (0.4444)
	Number of Observations	2308	1287	891	73	57

TABLE 5.2

REGRESSION RESULTS OF SEMI-LOG EARNINGS FUNCTION[a,b]

Explanatory Variable	Total	White Male	White Female	Black Male	Black Female
Schooling	0.0822 (20.78)	0.0765 (17.24)	0.0874 (10.48)	0.0743 (3.24)	0.1515 (6.07)
Reading	0.0484 (4.68)	0.0590 (4.67)	0.0339 (1.68)	0.0261 (0.73)	-0.0011 (-0.02)
Experience	0.0293 (8.40)	0.0474 (11.32)	0.0079 (1.22)	0.0178 (1.06)	0.0033 (0.14)
Experience-squared	-0.0005 (-6.20)	-0.0008 (-8.84)	-0.0001 (-0.44)	-0.0003 (-1.05)	-0.0001 (-0.13)
Father's Education	0.0004 (0.13)	0.0005 (0.14)	0.0012 (0.22)	0.0130 (0.67)	0.0143 (0.76)
Mother's Education	0.0031 (0.98)	0.0045 (1.22)	0.0041 (0.67)	-0.0092 (-0.45)	-0.0519 (-2.41)
Work Hours	0.8642 (42.33)	0.7529 (18.78)	0.8973 (33.23)	0.8761 (6.91)	0.7176 (6.15)
Race	-0.1560 (-4.02)	---	---	---	---
Sex	-0.6198 (-32.37)	---	---	---	---
Constant	-0.7384 (-9.48)	-0.6544 (-6.06)	-1.2656 (-9.18)	-0.7556 (-1.90)	-1.3414 (-2.88)
R^2	0.6093	0.3528	0.5055	0.5216	0.5210
F	577.14	143.72	187.18	15.36	12.77

[a] The dependent variable is the log of the number of thousands of dollars of annual earnings.

[b] t-values are expressed in parentheses below parameter estimates.

TABLE 5.3

REGRESSION RESULTS OF LINEAR EARNINGS FUNCTION [a,b]

Explanatory Variable	Total	White Male	White Female	Black Male	Black Female
Schooling	0.8087 (23.63)	0.9321 (19.22)	0.5550 (11.39)	0.6889 (4.10)	0.6580 (5.90)
Reading	0.2903 (3.24)	0.4699 (3.39)	0.0311 (0.26)	0.1274 (0.49)	-0.0824 (-0.37)
Experience	0.3651 (12.08)	0.5455 (11.87)	0.0891 (2.35)	0.0960 (0.78)	0.0409 (0.39)
Experience-squared	-0.0058 (-9.19)	-0.0088 (-9.08)	-0.0011 (-1.37)	-0.0010 (-0.45)	-0.0004 (-1.89)
Father's Education	0.0263 (1.07)	0.0193 (0.53)	0.0365 (1.17)	0.1585 (1.12)	0.0811 (0.97)
Mother's Education	-0.0175 (-0.64)	-0.0064 (-0.16)	0.0016 (0.05)	-0.1254 (-0.84)	-0.1671 (-1.74)
Work Hours	3.7545 (21.25)	5.3516 (12.18)	3.3272 (21.09)	3.6071 (3.88)	2.1396 (4.10)
Race	-1.7205 (-5.12)	---	---	---	---
Sex	-5.3440 (-32.24)	---	---	---	---
Constant	-9.7003 (-14.39)	-16.5161 (-13.95)	-8.2620 (-10.25)	-8.1316 (-2.79)	-6.5195 (3.13)
R^2	0.4974	0.3123	0.3400	0.4218	0.4395
F	366.20	119.73	94.35	10.28	9.21

[a]The dependent variable is the number of thousands of dollars of annual earnings.

[b]t-values are expressed in parentheses below parameter estimates.

TABLE 5.4

REGRESSION RESULTS OF WAGE DETERMINATION FUNCTION [a,b]

Explanatory Variable	Total	White Male	White Female	Black Male	Black Female
Schooling	0.5678 (16.76)	0.6724 (14.81)	0.3306 (5.55)	0.3919 (4.23)	0.4535 (7.16)
Reading	0.1025 (1.16)	0.1813 (1.40)	0.0459 (0.32)	0.0355 (0.25)	-0.0521 (-0.42)
Experience	0.2250 (7.53)	0.3058 (7.16)	0.1198 (2.58)	0.0679 (1.00)	-0.0031 (-0.05)
Experience-Squared	-0.0038 (-6.07)	-0.0051 (-5.68)	-0.0020 (-2.08)	-0.0009 (-0.70)	0.0002 (0.14)
Father's Education	-0.0248 (-1.02)	-0.0655 (-1.94)	0.0425 (1.12)	0.0580 (0.76)	0.0353 (0.75)
Mother's Education	-0.0588 (-2.15)	-0.1019 (-2.68)	0.0457 (1.05)	-0.0730 (-0.89)	-0.1520 (-2.79)
Race	-1.0865 (-3.27)	---	---	---	---
Sex	-2.6187 (-16.97)	---	---	---	---
Constant	-2.4999 (-4.21)	-4.0333 (-4.85)	-2.4904 (-2.64)	-0.9569 (-0.67)	-1.5405 (-1.34)
R^2	0.1794	0.1454	0.0529	0.2804	0.4397
F	91.07	52.37	11.95	6.47	10.88

[a]The dependent variable is wages expressed in dollars per hour.

[b]t-values are expressed in parentheses below parameter estimates.

TABLE 5.5

REGRESSION RESULTS OF WORK HOURS FUNCTION [a,b]

Explanatory Variable	Total	White Male	White Female	Black Male	Black Female
Schooling	0.0152 (4.39)	0.0063 (2.39)	0.0246 (2.86)	-0.0018 (-0.09)	0.0404 (1.29)
Reading	0.0001 (0.01)	0.0058 (0.84)	-0.0129 (-0.64)	0.0289 (1.01)	-0.0833 (-1.84)
Experience	0.0139 (4.76)	0.0178 (7.64)	0.0051 (0.77)	0.0020 (0.15)	0.0265 (1.21)
Experience-squared	-0.0002 (-3.14)	-0.0003 (-6.42)	0 (0.22)	0.0001 (0.23)	-0.0005 (-1.18)
Father's Education	-0.0008 (-0.33)	0.0018 (0.99)	-0.0071 (-1.33)	0.0306 (1.98)	0.0308 (1.72)
Mother's Education	0.0024 (0.90)	0.0011 (0.52)	0.0029 (0.47)	0.0071 (0.44)	0.0049 (0.22)
Other Income	-0.0011 (-1.62)	0.0014 (2.60)	-0.0034 (-2.55)	0.0065 (1.14)	-0.0006 (-0.10)
Wage	-0.0216 (-12.81)	-0.0178 (-13.96)	-0.0306 (-7.77)	-0.0082 (-0.41)	-0.0653 (-1.57)
Race	0.0044 (0.14)	---	---	---	---
Sex	-0.3704 (-23.15)	---	---	---	---
Constant	1.6779 (28.86)	1.7342 (38.01)	1.3547 (10.10)	1.5319 (5.25)	0.8218 (1.87)
R^2	0.1648	0.1172	0.0645	0.1498	0.162
F	65.68	30.60	11.05	2.15	1.72

[a]The dependent variable is the number of thousands of hours the respondent worked in 1972.

[b]t-values are expressed in parentheses below parameter estimates.

TABLE 5.6

REGRESSION RESULTS OF DETERMINANTS OF SCHOOLING [a,b]

Explanatory Variable	Total	White Male	White Female	Black Male	Black Female
Father's Education	0.1776 (13.29)	0.1903 (10.04)	0.1724 (9.21)	0.1769 (1.87)	0.0573 (0.72)
Mother's Education	0.1856 (12.28)	0.1488 (6.94)	0.2168 (10.19)	0.3811 (3.93)	0.2715 (3.00)
Age	-0.0149 (-3.38)	-0.0186 (-2.86)	-0.0078 (-1.32)	-0.0868 (-3.89)	0.0309 (1.18)
Race	-0.8637 (-4.71)	---	---	---	---
Sex	-0.2570 (-2.94)	---	---	---	---
Constant	10.1987 (42.38)	10.6111 (30.50)	9.3639 (28.35)	9.7532 (8.69)	8.2909 (5.98)
R^2	0.2402	0.1947	0.3090	0.4640	0.1883
F	210.80	149.09	191.65	29.62	6.66

[a]The dependent variable is the number of years of schooling attained by the respondent.

[b]t-values are expressed in parentheses below parameter estimates.

TABLE 5.7

REGRESSION RESULTS OF DETERMINANTS OF READING [a,b]

Explanatory Variable	Total	White Male	White Female	Black Male	Black Female
Father's Education	0.0001 (0.02)	-0.0037 (-0.60)	-0.0122 (-1.65)	0.1134 (2.20)	0.1612 (4.35)
Mother's Education	0.0322 (6.03)	0.0430 (6.36)	0.0270 (3.21)	-0.0527 (-0.95)	-0.1367 (-3.10)
Schooling	0.1331 (22.24)	0.1274 (17.60)	0.1382 (12.99)	0.2142 (4.06)	0.2070 (4.14)
Age	-0.0029 (-1.92)	-0.0026 (-1.27)	-0.0051 (-2.24)	0.0227 (1.78)	0.0004 (0.03)
Race	-0.8759 (-13.79)	---	---	---	---
Sex	0.0517 (1.71)	---	---	---	---
Constant	-1.8590 (-18.02)	-1.8639 (-14.04)	-1.6296 (-10.15)	-4.8244 (-6.10)	-3.4641 (-4.53)
R^2	0.2637	0.2389	0.1979	0.2859	0.3330
F	199.02	145.12	79.28	10.17	10.63

[a]The dependent variable is the respondent's standardized reading score on a literacy test.

[b]t-values are expressed in parentheses below the parameter estimates.

TABLE 5.8

ALTERNATIVE MEASURES OF BENEFIT OF SCHOOLING

	Total	White Male	White Female	Black Male	Black Female
I. Marginal Benefit[a]					
Partial Benefit	$ 1,042	$ 1,278	$ 581	$ 706	$ 863
Full Benefit	931	1,121	517	663	783
Wage Effect	939	1,186	482	687	805
Hour Effect	-8	-65	35	-23	-22
Adjusted Benefit	1,046	1,229	602	743	916
Present value of Full Benefit					
No discount	46,236	56,306	25,520	33,188	36,116
5% discount	15,594	18,518	8,871	12,331	13,586
10% discount	7,972	9,270	4,655	6,809	7,447
II. Total Benefit					
Compulsory High School Education					
Representative Individual[b]	$ 2,900	$ 3,810	$1,370	$2,580	$1,940
National Program (Billions)[c]	142.3 (120.9)	76.0	29.9	8.1	6.9
Compulsory College Education					
Representative Individual[b]	4,040	4,910	2,200	3,430	3,380
National Program[c] (Billions)	428.5 (332.8)	194.1	103.7	15.1	19.9

[a]These are marginal private benefits of one additional year of school.

[b]The total benefits given for a representative individual are computed by $\int_0^m (dY/dX_1)\ f(X_1|X_1 < m)\ dX_1$ where $f(X_1/X_1 < m)$ is the conditional distribution of individuals by educational level, and m is either 12 or 16, corresponding to high school and college graduation respectively.

[c]The national program benefits are computed by multiplying the representative individual benefits by the corresponding total numbers of individuals (age 25 and over) completing less than high school or college education for 1970 taken from Statistical Abstract of the United States 1972, No. 168, p. 112. Figures in parentheses for total samples are derived by summing the benefits for all four subsamples.

TABLE 5.9

ALTERNATIVE MEASURES OF BENEFITS OF READING COMPETENCE

		White		Black	
	Total	Male	Female	Male	Female
I. Marginal Benefit					
Partial Benefit	$ 172	$ 375	$ 12	$ 187	$ -341
Full Benefit	172	375	12	187	-341
Wage Effect	185	356	72	66	-95
Hour Effect	-13	19	-60	121	-246
Adjusted Benefit					
Present value of Full Benefit	$ 206	$ 369	$ 90	$ 71	$ -108
No discount	8,267	17,606	803	8,956	-16,357
5% discount	3,224	6,830	384	3,385	-6,359
10% discount	1,832	3,848	249	1,876	-3,604
II. Total Benefit					
Low Reading Target (Standardized reading score = -1)					
Representative Individual[a]	$ 161	$ 323	$ 12	$ 273	$ -392
National Program (Millions)[b]	2,552	2,002	71	609	-873
	(1,809)				
High Reading Target (Standardized reading score = 0)					
Representative Individual[a]	163	332	12	340	-490
National Program (Millions)[b]	7,216	6,004	237	1,098	-1,863
	(5,476)				

[a] The total benefits given for a representative individual are computed by

$$\int_{-\infty}^{m} (dY/dX_2)\, f(X_2 | X_2 < m)\, dX_2 \quad \text{where} \quad f(X_2/X_2 < m) \text{ is the conditional distribution}$$

of individuals by standardized reading score, and m is either -1 or zero, corresponding to a low or high target compensatory reading program.

[b] The national program benefits are computed by multiplying the representative individual benefits by the corresponding estimated numbers of individuals whose reading competence, measured by standardized scores, are below -1 or zero, using the frequency distributions of 1972 National Reading Survey and population figures (age 25 and over) for 1970 obtained from Statistical Abstract of the United States 1972, No. 168, p. 112. Figures in parentheses for total sample are derived by summing the benefits for all four subsamples.

APPENDIX A: CORRELATION MATRICES

Tables A-1 through A-5 contain the correlation matrices of the variables used in the study. Table A-1 is the correlation matrix for the total sample; Tables A-2 to A-4 are the matrices for the four subsamples.

APPENDIX TABLE A-1

CORRELATION MATRIX: TOTAL SAMPLE

	Y Earnings	Ln Y Log of Earnings	Y_1 Wage Rate	Y_2 Work Hours	X_1 Schooling	X_2 Reading	X_3 Experience	X_3^2 Experience Squared	X_4 Father's Education	X_5 Mother's Education	X_6 Other Income	X_7 Age	X_8 Race
Ln Y Log of Earnings	0.898												
Y_1 Wage Rate	0.620	0.508											
Y_2 Work Hours	0.430	0.620	-0.092										
X_1 Schooling	0.340	0.287	0.281	-0.007									
X_2 Reading	0.195	0.175	0.147	-0.016	0.456								
X_3 Experience	0.003	-0.013	-0.019	0.063	-0.426	-0.234							
X_3^2 Experience-squared	-0.023	-0.029	-0.040	0.057	-0.411	-0.246	0.973						
X_4 Father's Education	0.105	0.091	0.051	-0.023	0.437	0.263	-0.322	-0.295					
X_5 Mother's Education	0.108	0.113	0.047	-0.004	0.434	0.302	-0.377	-0.361	0.643				
X_6 Other Income	0.119	0.013	0.135	-0.111	0.189	0.100	0.045	0.025	0.106	0.065			
X_7 Age	0.098	0.066	0.057	0.067	-0.185	-0.128	0.968	0.943	-0.228	-0.290	0.101		
X_8 Race	-0.122	-0.094	-0.095	0.006	-0.092	-0.243	-0.004	-0.001	-0.052	-0.031	-0.066	-0.030	
X_9 Sex	-0.522	-0.545	-0.276	-0.336	-0.042	0.004	0.048	0.044	0.027	-0.002	0.204	0.041	0.023

APPENDIX TABLE A-2

CORRELATION MATRIX: WHITE MALE

	Y Earnings	Ln Y Log of Earnings	Y_1 Wage Rate	Y_2 Work Hours	X_1 Schooling	X_2 Reading	X_3 Experience	X_3^2 Experience Squared	X_4 Father's Education	X_5 Mother's Education	X_6 Other Income
Ln Y Log of Earnings	0.928										
Y_1 Wage Rate	0.613	0.545									
Y_2 Work Hours	0.256	0.373	-0.274								
X_1 Schooling	0.380	0.349	0.297	-0.051							
X_2 Reading	0.241	0.254	0.151	-0.004	0.462						
X_3 Experience	0.062	0.048	0.021	0.091	-0.437	-0.258					
X_3^2 Experience-squared	0.010	-0.004	-0.013	0.066	-0.428	-0.273	0.972				
X_4 Father's Education	0.125	0.133	0.006	0.018	0.407	0.261	-0.321	0.292			
X_5 Mother's Education	0.107	0.136	-0.012	0.017	0.382	0.322	-0.370	-0.353	0.635		
X_6 Other Income	0.358	0.256	0.266	0.007	0.209	0.070	0.071	0.059	0.098	0.062	
X_7 Age	-.180	0.156	0.111	0.085	-0.181	-0.145	0.965	0.936	-2.230	-0.291	0.140

APPENDIX TABLE A-3

CORRELATION MATRIX: WHITE FEMALE

	Y Earnings	Ln Y Log of Earnings	Y_1 Wage Rate	Y_2 Work Hours	X_1 Schooling	X_2 Reading	X_3 Experience	X_3^2 Experience Squared	X_4 Father's Education	X_5 Mother's Education	X_6 Other Income
Ln Y Log of Earnings	0.892										
Y_1 Wage Rate	0.476	0.336									
Y_2 Work Hours	0.480	0.654	-0.212								
X_1 Schooling	0.303	0.252	0.209	-0.026							
X_2 Reading	0.119	0.118	0.102	-0.040	0.432						
[illegible] Experience	0.003	0.008	-0.040	0.102	-0.408	-0.243					
X_3^2 Experience-squared	0.005	0.015	-0.048	0.106	-0.381	-0.249	0.974				
X_4 Father's Education	0.137	0.092	0.139	-0.068	0.487	0.[illegible]25	-0.332	-0.303			
X_5 Mother's Education	0.132	0.110	0.140	-0.051	0.512	0.293	-0.399	-0.378	0.652		
X_6 Other Income	0.108	0.040	0.092	-0.081	0.173	0.105	-0.007	-0.038	0.108	0.082	
X_7 Age	0.077	0.070	0.008	0.103	-0.195	-0.156	0.975	0.954	0.082	-0.304	0.034

APPENDIX TABLE A-4

CORRELATION MATRIX: BLACK MALE

	Y Earnings	Ln Y Log of Earnings	Y_1 Wage Rate	Y_2 Work Hours	X_1 Schooling	X_2 Reading	X_3 Experience	X_3^2 Experience Squared	X_4 Father's Education	X_5 Mother's Education	X_6 Other Income
Ln Y Log of Earnings	0.914										
Y_1 Wage Rate	0.907	0.807									
Y_2 Work Hours	0.429	0.597	0.084								
X_1 Schooling	0.513	0.464	0.506	0.131							
X_2 Reading	0.361	0.360	0.291	0.210	0.466						
X_3 Experience	-0.151	-0.157	-0.205	0.094	-0.592	-0.127					
X_3^2 Experience-squared	-0.163	-0.181	-0.215	0.078	-0.579	-0.145	0.974				
X_4 Father's Education	0.427	0.427	0.312	0.317	0.513	0.409	-0.189	-0.227			
X_5 Mother's Education	0.365	0.379	0.289	0.237	0.611	0.323	-0.323	-0.357	0.719		
X_6 Other Income	0.294	0.236	0.262	0.097	0.217	0.012	0.029	0.010	-0.052	0.061	
X_7 Age	-0.018	-0.041	-0.083	0.148	-0.378	-0.006	0.970	0.944	-0.062	-0.187	0.099

APPENDIX TABLE A-5

CORRELATION MATRIX: BLACK FEMALE

	Y Earnings	Ln Y Log of Earnings	Y Wage Rate	Y Work Hours	X Schooling	X Reading	X Experience	X Experience Squared	X Father's Education	X Mother's Education	X Other Income
Ln Y Log of Earnings	0.929										
Y_1 Wage Rate	0.821	0.724									
Y_2 Work Hours	0.368	0.483	-0.156								
X_1 Schooling	0.547	0.524	0.602	0.063							
X_2 Reading	0.186	0.193	0.244	-0.084	0.408						
X_3 Experience	-0.075	-0.131	-0.070	-0.088	-0.240	-0.214					
X_3^2 Experience-squared	-0.060	-0.117	-0.032	-0.113	-0.203	-0.212	0.975				
X_4 Father's Education	0.203	0.203	0.060	0.244	0.296	0.410	-0.445	-0.424			
X_5 Mother's Education	0.195	0.167	0.005	0.289	0.417	0.137	-0.314	-0.335	0.646		
X_6 Other Income	0.427	0.339	0.511	-0.085	0.436	0.277	-0.034	-0.005	0.106	-0.124	

APPENDIX B

TESTS FOR LINEARITY OF EFFECTS BY USING DISCRETE VARIABLES

The purpose of this appendix is to report some supplementary regression results of empirical earnings functions using dummy variables to represent different levels of schooling and reading competence. In Section II, schooling and reading competence are represented by years of schooling and standardized scores in computing empirical earnings function, implicitly assuming that the effects of these two variables (measured by the associated coefficients) are the same at different levels of schooling and reading competence. This appendix summarizes the empirical results of an attempt to verify the reasonability of this implicit assumption. Specifically, the schooling and reading scores variables in eqs. (1) and (2) are substituted by the following set of six dummy variables, four of them representing schooling and the other two representing reading competence:

School Dummy 1: 5 to 8 years of schooling = 1, otherwise = 0

School Dummy 2: 9 to 12 years of schooling = 1, otherwise = 0

School Dummy 3: 13 to 16 years of schooling = 1, otherwise = 0

School Dummy 4: 16 and more years of schooling = 1, otherwise = 0

Reading Dummy L: Standardized scores below minus one = 1, otherwise = 0

Reading Dummy H: Standardized scores above one = 1, otherwise = 0.

The results of the empirical earnings function using these dummy variables for schooling and reading competence are reported in Appendix Tables B-1 and B-2, using logarithmic values of earnings and earnings respectively as the dependent variables. These results are comparable with those reported in Tables 2 and 3 in the text. In general, the results shown here are similar to the corresponding results shown in Tables 2 and 3 of the text. It is important to note that while the assumption of constant schooling effect in eqs. (1) and (2) appears to be acceptable, the similar assumption for reading competence seems to be more questionable. This conclusion is derived from the observation that the estimated coefficients of the four school dummies seem to increase at a roughly contant rate as schooling level increases, and that the magnitudes of the estimated coefficients of the two reading dummies (in terms of absolute values) are considerably different from each other. The results of Appendix Tables B-1 and B-2 generally show that individuals with low reading scores can be expected to have lower earnings, with a possible exception of black females. Whether individuals with high reading scores can be expected to have higher earnings, however, is not very conclusive from our results, perhaps because the reading test instruments were designed only to reveal functional reading ability. This result suggests that the economic benefit of reading competence may tend to underestimate the true effect. Finally, it may be pointed out that while the school dummies are usually statistically significant, especially for school dummies 3 and 4, school dummies 1 and 2 for white females turn out to have negative estimated coefficients.

APPENDIX TABLE B-1

SUPPLEMENTARY SEMI-LOG EARNINGS FUNCTION[a,b]

Explanatory Variable	Total	White		Black	
		Male	Female	Male	Female
School Dummy 1	0.3402 (2.80)	0.7414 (4.26)	-0.3494 (-1.33)	0.2603 (0.96)	0.1762 (0.31)
School Dummy 2	0.5585 (4.63)	0.9500 (5.46)	-0.1277 (-0.49)	0.4654 (1.76)	1.0214 (1.90)
School Dummy 3	0.8271 (6.75)	1.2156 (6.93)	0.1007 (0.38)	0.7528 (2.46)	1.5708 (2.88)
School Dummy 4	1.0789 (8.64)	1.4070 (7.93)	0.4815 (1.80)	0.9581 (2.50)	2.0418 (3.36)
Reading Dummy L	-0.1634 (-5.62)	-0.1667 (-4.58)	-0.1709 (-3.12)	-0.1403 (-1.28)	0.0820 (0.58)
Reading Dummy H	0.0360 (1.37)	0.0432 (1.44)	0.0250 (0.51)	0.0075 (0.03)	-0.0420 (-0.19)
Experience	0.0295 (8.04)	0.0462 (10.49)	0.0117 (1.73)	0.0117 (0.66)	-0.0080 (-0.33)
Experience-squared	-0.0005 (-6.1413)	-0.0008 (-8.30)	-0.0002 (-1.11)	-0.0002 (-0.68)	0.0002 (0.32)
Work Hours	0.8715 (42.02)	0.7637 (18.65)	0.9051 (33.31)	0.8574 (6.13)	0.6969 (5.45)
Father's Education	0.0021 (0.73)	0.0028 (0.83)	0.0025 (0.47)	0.0209 (1.02)	0.0144 (0.73)
Mother's Education	0.0216 (1.44)	0.0046 (1.23)	0.0078 (1.28)	-0.0051 -0.25	-0.0511 (-2.35)
Race	-0.1536 (-3.91)	---	---	---	---
Sex	-0.6131 (-31.34)	---	---	---	---
R^2	0.5990	0.3281	0.5043	0.5199	0.5181
F	382.31	81.78	118.19	9.32	7.64

a. The dependent variable is the log of the number of thousands of dollars of annual earnings.

b. t-values are expressed in parentheses below parameter estimates.

APPENDIX TABLE B-2

SUPPLEMENTARY LINEAR EARNINGS FUNCTION [a,b]

Explanatory Variable	Total	White Male	White Female	Black Male	Black Female
School Dummy 1	0.6504 (0.62)	2.7627 (1.44)	-2.0170 (-1.32)	0.7079 (0.37)	0.2265 (0.09)
School Dummy 2	2.6749 (2.55)	4.9729 (2.60)	-0.5910 (-0.39)	2.0993 (1.12)	2.9907 (1.28)
School Dummy 3	5.3913 (5.07)	8.4228 (4.37)	0.7696 (0.50)	6.6739 (3.07)	5.9279 (2.50)
School Dummy 4	7.9786 (7.36)	10.9533 (5.62)	3.4436 (2.21)	6.0697 (2.22)	7.73 (2.92)
Reading Dummy L	-1.1352 (-4.50)	-1.4766 (-3.69)	-0.7607 (-2.39)	-0.6632 (-0.85)	0.9217 (1.49)
Reading Dummy H	0.1523 (0.67)	0.3677 (1.12)	-0.2148 (-0.75)	-1.7118 (-0.95)	0.2454 (0.26)
Experience	0.3850 (12.10)	0.5641 (11.65)	0.1156 (2.93)	0.1086 (0.86)	-0.0146 (-0.14)
Experience-squared	-0.0065 (-9.62)	-0.0095 (-9.24)	-0.0017 (-2.07)	-0.0017 (-0.71)	0.0006 (0.26)
Work Hours	3.8571 (21.42)	5.4476 (12.10)	3.3853 (21.40)	3.7299 (3.75)	2.0408 (3.67)
Father's Education	0.0421 (1.70)	0.0429 (1.16)	0.0437 (1.401)	0.2102 (1.44)	0.0867 (1.01)
Mother's Education	0.0002 (0.01)	-0.0006 (-0.02)	0.0250 (0.70)	-0.0605 (-0.42)	-0.1521 (-1.60)
Race	-1.7440 (-5.12)	---	---	---	---
Sex	-5.2478 (-30.91)	---	---	---	---
R^2	0.4812	0.2825	0.3427	0.4529	0.4660
F	237.35	65.94	60.56	7.12	6.20

a. The dependent variable is the number of thousands of dollars of annual earnings.

b. t-values are expressed in parentheses below parameter estimates.

APPENDIX C

SIMULTANEOUS LABOR MARKET MODEL

The purpose of this appendix is to supplement the analysis of Section III by considering a simultaneous model of labor market relationships. The simultaneous model in this appendix is different from the recursive model reported in the text mainly in allowing work hours to affect wage rate in wage determination function represented by eq. (3), i.e., $\gamma \neq 0$. In addition, for identification purpose, a new variable referred to as employment status is introduced into eq. (3) as an additional explanatory variable. Thus eq. (3) may be regarded as the inverse demand function for labor, and eq. (4) the supply function of labor. Because both wage rate and work hours are endogenous in the simultaneous model, some simultaneous equation approach must be considered for estimating the parameters in eqs. (3) and (4). In this appendix, the results of the Two-Stage Least Squares (TSLS) procedure are presented in Appendix Tables C-1 and C-2 for the wage determination and work hours functions respectively. The results of the Ordinary Least Squares (OLS) procedure for the wage determination function, not reported here, are in general very similar to those of the TSLS procedure presented in Appendix Table C-1. The results of the OLS procedure for the work hours function are identical with those shown in Table 5 in the text.

The results reported in Appendix Table C-1 are obtained by restricting the coefficients associated with employment status to the corresponding estimates obtained in the first stage. These additional restrictions were introduced because the predicted work hours were so highly correlated with employment status (full-time salaried workers or not) that the usual second stage computation in the TSLS procedure became infeasible due to singularity of a matrix to be inverted. In Appendix Table C-2, the coefficient of schooling for black males was restricted to be zero, since the corresponding estimate in the first stage is negative, contrary to the usual expectation, and without such a restriction the computation was not feasible also due to singularity. Aside from these restrictions which were imposed only to avoid computational difficulties, the empirical results of Appendix Tables C-1 and C-2 are obtained following the usual TSLS procedure. One of the reasons the recursive model was selected for discussion in the text is based on the fact that, while work hours seem to be a significant factor determining wage rate for the whites, the same does not appear to be true for the blacks. In addition, as was just mentioned, the introduction of an employment status variable for identification purpose resulted in computational difficulties. We thus present these simultaneous estimations only tentatively.

The empirical results of Appendix Table C-1 are similar to those of Table 4, which did not include employment status and predicted work hours as additional explanatory variables. The empirical results of Appendix Table C-2, however, are considerably different from those of Table 5 in the text. Recall that, besides a minor restriction introduced on the parameter associated with schooling variable for black males, the only difference between Appendix Table C-2 and Table 5 is the use of observed wage rate or predicted wage rate as an explanatory variable. As Appendix Table C-2 shows, the use of predicted wage rate has in general increased the goodness-of-fit considerably, except for black males. In fact, all coefficients in the work hours functions for white males, white females and black females are highly statistically significant, except work experience for black females. In general, the effects of schooling, reading and work experience all appear to be much stronger than those of the OLS estimates employed for the recursive model as shown in Table 5. In particular, the effects of reading and work hours are highly significantly positive for white males and white females, but negative for black females. The estimated effects of father's education and mother's education on work hours are very different from those of the OLS estimates.

APPENDIX TABLE C-1

REGRESSION RESULTS OF WAGE DETERMINATION FUNCTION [a] (TSLS)

Explanatory Variable	Total	White		Black	
		Male	Female	Male	Female
Schooling	0.5758 (17.08)	0.6545 (14.45)	0.3770 (6.34)	0.3696 (3.97)	0.4559 (7.24)
Reading	0.0995 (1.13)	0.1909 (1.48)	0.0353 (0.25)	0.0519 (0.36)	-0.0872 (-0.70)
Experience	0.2461 (8.25)	0.3648 (8.34)	0.1272 (2.75)	0.0671 (0.98)	0.0153 (0.26)
Experience-Squared	-0.0041 (-6.53)	-0.0061 (-6.71)	-0.0019 (-1.97)	-0.0009 (-0.71)	-0.0002 (-0.16)
Father's Education	-0.0248 (-1.03)	-0.0500 (-1.48	0.0235 (0.62)	0.0641 (0.81)	0.0506 (1.06)
Mother's Education	-0.0505 (-1.85)	-0.0883 (-2.33)	0.0425 (0.98)	-0.0704 (-0.85)	-0.1439 (-2.65)
Employment Status	-0.6399 (b)	-0.2392 (b)	-1.5554 (b)	-0.7422 (b)	0.1926 (b)
Work Hours	-2.5813 (-9.76)	-4.6021 (-5.70)	-2.4739 (-9.91)	-0.7307 (-1.25)	-0.3364 (-0.95)
Race	-1.0635 (-3.21)	---	---	---	---
Sex	-3.3083 (-18.88)	---	---	---	---
Constant	2.1277 (2.84)	4.2151 (2.54)	1.5366 (1.52)	0.7638 (0.46)	-1.4445 (-1.21)
R^2	0.1953	0.1597	0.1225	0.2630	0.4461
F	89.83	50.11	25.56	5.03	9.45

[a] t-values are placed in parentheses below the parameter estimates.

[b] These estimates were restricted to the given value to avoid multicollinearity.

APPENDIX TABLE C-2.

REGRESSION RESULTS OF WORK HOURS FUNCTION [a] (TSLS)

Explanatory Variable		White		Black	
	Total	Male	Female	Male	Female
Schooling	0.3680 (47.91)	0.1475 (23.01)	0.4290 (38.59)	(b)	0.7554 (13.43)
Reading	0.0718 (10.60)	0.0573 (8.55)	0.0375 (2.81)	0.0298 (0.99)	-0.1750 (-6.68)
Experience	0.1493 (41.83)	0.0839 (23.99)	0.1421 (26.19)	0.0022 (0.17)	0.0011 (0.09)
Experience-Squared	-0.0025 (-37.49)	-0.0015 (-22.28)	-0.0022 (-21.06)	0.0001 (0.23)	0.0002 (0.67)
Father's Education	0.0237 (-12.70)	-0.0171 (-9.33)	0.0433 (11.66)	0.0314 (1.84)	0.0443 (4.41)
Mother's Education	-0.0355 (-16.29)	-0.0240 (-11.26)	0.0628 (14.73)	0.0067 (0.44)	-0.1788 (-9.59)
Other Income	0.0336 (38.76)	0.0217 (21.98)	0.0155 (15.84)	0.0069 (0.96)	0.0642 (11.05)
Wage (predicted)	-0.7055 (-50.30)	-0.2611 (-25.29)	-1.3070 (-43.02)	-0.0156 (0.26)	-2.0021 (-13.65)
Race	-0.6599 (-23.31)	---	---	---	---
Sex	-2.3348 (-55.63)	---	---	---	---
Constant	0.4732 (9.24)	1.0731 (21.52)	-1.6306 (-14.46)	1.5378 (5.49)	-0.8942 (-3.23)
R^2	0.5020	0.2751	0.5994	0.1484	0.7381
F	335.65	87.54	239.55	2.46	28.60

[a] t-values are placed in parentheses below the parameter estimates.

[b] This parameter was restricted to be zero.

APPENDIX D

SIMULTANEOUS MODEL OF EDUCATIONAL ACHIEVEMENTS

In Section IV of the text, the empirical results of a recursive model of educational achievements have been examined. The purpose of this Appendix is to supplement those results by examining the empirical results of a simultaneous model in which not only schooling is assumed to affect reading but also reading is assumed to affect schooling. In order to identify eqs. (5) and (6), in eq. (5) we replaced father's education and mother's education by a single variable constructed by summing the years of schooling of both parents, i.e., restricting the parameters associated with father's education and mother's education to be the same. In addition, in eq. (6), we assumed that father's education does not affect reading for total sample and subsamples of white males and white females but that mother's education does. For black males and females, however, we assumed father's education, rather than mother's education, affects reading. These restrictions, imposed for identification purposes, are largely based on empirical results and are admittedly somewhat arbitrary.

The empirical results of the Two Stage Least Squares (TSLS) estimation of the simultaneous model of educational achivements are presented in Appendix Tables D-1 and D-2 for the determinants of schooling and reading respectively. The corresponding results of the Ordinary Least Squares (OLS) estimation are not reported here. The most striking differences between these two sets of estimated values using TSLS and OLS respectively are the coefficients associated with schooling in eq. (6), representing the interaction between these two variables. While these coefficients are highly significant when the observed reading scores and schooling level were used in eqs. (5) and (6), as in the OLS procedure, the same is not generally true when the predicted reading scores and schooling level were used instead, as in the TSLS procedure (Appendix Tables D-1 and D-2).

Among the determinants of schooling examined in Appendix Table D-1, father's and mother's education clearly has a significantly positive effect for all subsamples. The effect of age is generally negative, reflecting a general trend of increasing educational level, with a possible exception of black females. The effect of reading on schooling is generally not significant and mostly turns out to be negative, possibly because of the oversimplified structure of our simultaneous model. Among the determinants of reading scores considered in Appendix Table D-2, mother's education is statistically significant for whites and father's education is for blacks. The effect of age on reading is negative for whites and positive for blacks, but statistically significant only for white females. The effect of schooling on reading is generally positive, except for black females. However, only the estimated coefficients for the total sample and the subsample of white males are statistically significant. In general, our results do indicate that the recursive model considered in the text ($\lambda=0$, $\mu=0$) is more plausible than the other type of recursive model ($\lambda\neq0$, $\mu=0$). Finally, the effects of race and sex are obvious either from the results of dummy variables or from comparisons among corresponding regressions obtained for various subsamples.

APPENDIX TABLE D-1

REGRESSION RESULTS OF DETERMINANTS OF SCHOOLING[a] (TSLS)

Explanatory Variable	Total	White Male	White Female	Black Male	Black Female
Father's Education[b]	0.1719 (5.59)	0.2109 (5.87)	0.1610 (5.91)	0.4294 (3.14)	0.2035 (4.07)
Mother's Education[b]	0.1719 (5.59)	0.2109 (5.87)	0.1610 (5.91)	0.4294 (3.14)	0.2035 (4.07)
Reading (predicted)	0.2412 (0.31)	-1.0032 (-1.14)	0.9781 (1.23)	-1.6703 (-1.45)	-0.8449 (-1.40)
Age	-0.0137 (-2.27)	-0.0235 (-2.91)	-0.0018 (-0.23)	-0.0799 (-3.33)	0.0366 (1.42)
Race	-0.6247 (-0.80)	---	---	---	---
Sex	-0.2613 (-2.96)	---	---	---	---
Constant	10.3198 (24.33)	10.0978 (19.26)	9.6922 (24.75)	5.1837 (1.19)	6.8140 (3.94)
R^2	0.2402	0.1947	0.3090	0.4640	0.1883
F	210.80	149.09	191.65	29.62	6.66

[a] t-values are placed in parentheses below the parameter estimates.

[b] These estimates were constrained to have the same magnitude for father's education as for mother's education.

APPENDIX TABLE D-2

REGRESSION RESULTS OF DETERMINANTS OF READING [a] (TSLS)

Explanatory Variable	White			Black	
	Total	Male	Female	Male	Female
Father's Education [b]	---	---	---	0.1378 (1.83)	0.1900 (4.05)
Mother's Education [c]	0.0321 (3.32)	0.0458 (4.17)	0.0424 (2.58)	---	---
Schooling (predicted)	0.1335 (4.79)	0.1082 (3.23)	0.0675 (1.53)	0.0758 (0.52)	-0.2963 (-1.76)
Age	-0.0029 (-1.70)	-0.0029 (-1.27)	-0.0056 (-2.30)	0.0107 (0.54)	0.0159 (1.10)
Race	-0.8755 (11.99)	---	---	---	---
Sex	0.0518 (1.57)	---	---	---	---
Constant	-1.8635 (-5.96)	-1.6597 (-4.21)	-0.9674 (-2.12)	-3.4751 (-1.95)	0.7087 (0.41)
R^2	0.1545	0.1114	0.0925	0.1702	0.1986
F	121.87	77.26	43.71	7.02	7.12

[a] t-values are placed in parentheses below the parameter estimates.

[b] Father's education was constrained to have a zero coefficient for the total sample and the white subsamples.

[c] Mother's education was constrained to have a zero coefficient for the black subsamples.